

现代归纳逻辑与概率逻辑

科学出版社

现代归纳逻辑与概率逻辑

李 小 五 编





科 学 出 版 社

1 9 9 2

（京）新登字092号

内 容 简 介

本书系统地介绍了现代归纳逻辑和概率逻辑的理论，详细地分析了国外各重要的逻辑系统，以及它们之间的关系．主要内容包括：Keynes的概率逻辑、Reichenbach的概率逻辑与归纳逻辑、Carnap的归纳逻辑、归纳方法的多维连续统理论、条件句概率逻辑、模态归纳逻辑、概率语义学、无穷概率逻辑。

本书可供高等院校数学系、哲学系、计算机科学系的师生阅读，也可作为科研人员的参考书．

现代归纳逻辑与概率逻辑

李小五 编

责任编辑 林 鹏

科学出版社出版

北京东黄城根北街16号

邮政编码：100707

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1992年12月第一版 开本：787×1092 1/32

1992年12月第一次印刷 印张：8

印数：1—2 000 字数：178 000

ISBN 7-03-003024-9/O·557

定价：7.30元

序

对概然命题的思考至少可以追溯到亚里士多德．他曾说："概然就是经常发生的事物．"([1],p. 80).

归纳法作为一种方法论和推理模式，在古希腊就为人所意识到，这时的归纳法是指从特殊事例进行概括的方法，或者从个别到一般的推理．亚里士多德在《分析篇》、《论辩篇》中就谈到不完全归纳法和完全归纳法．在书中归纳法有三种：简单枚举法、完全归纳法、直觉归纳法．到了近代，英国哲学家 F. Bacon（1561—1626）提出了有别于枚举归纳法的消除归纳法．后人，例如 L.J. Cohen，对此给予很高的赞誉．但是 Bacon 的归纳法，就像 R. Carnap 所认为的那样，与其说是一种推理模式，还不如说是一种实验科学中经常使用的科学方法或程序．直到19世纪中叶，J. S. Mill 才为传统归纳推理建立了较完善的体系．在他的体系中，对归纳法做了最广泛的理解．例如，他认为演绎三段论也是一种归纳法，从而把演绎三段论也纳入他的体系．这主要反映在他的《逻辑体系》一书中．在书中，Mill 总结了前人在归纳法方面的成就，提出了著名的 Mill 五法，比较完善地建立了古典归纳逻辑．

I. Hacking 认为概率这个概念是从文艺复兴时期"意见"这个概念转化而来的（参见 I. Hacking: The Emergence of Probability, p.185). 1655—1665 年是对概率研究发展的时期．1657 年，C. Huygens 写了第一本公开出版的关于概率的教科书．大约在这个时期，B. Pascal 第一次把概率推理运

用于解决问题，从而建立了决策论（decision theory），1662年在 A. Arnauld 和 P. Nicole 著的《王港逻辑》(Port Royal Logic）一书的结尾已经对此做了概括.

G.W. Leibniz 是第一个明确提出要建立概率逻辑的人，虽然他没有建立概率逻辑的公理化系统，甚至在形式上也没有对此做出什么贡献，但他对概率的哲学本性做了许多讨论.

现代归纳逻辑是用数理逻辑和概率论等数学工具对归纳推理（我们称结论断定的范围超出前提断定的范围的推理为概然推理，其中一部分是常见的、为古典归纳逻辑所讨论过的称为归纳推理）进行数量化、形式化和公理化的研究. 现代归纳逻辑严格说不是一种统一的公认的逻辑系统，而是不同的、往往是互相冲突的归纳逻辑理论的总称. 现代归纳逻辑主要关心的是建立它的形式语义学(即归纳概率演算)，提出合理地求初始概率（即如何使求初始概率的方法能够刻划古典归纳推理的过程)的方法（这种方法通常称为语义解释）以及用这些工具处理古典归纳法.

另一方面，概率逻辑是用数理逻辑和概率论、测度论等数学工具对概然推理进行数量化、形式化和公理化的研究. 它可以分为有穷概率逻辑和无穷概率逻辑. 前者是用经典数理逻辑和概率论来研究概然推理，后者用一阶逻辑的无穷扩张，可容集合论、测度论和模型论来研究概然推理.

所以，应该说现代归纳逻辑和概率逻辑是两种不同的逻辑，它们的建立是出于不同的兴趣，服务于不同的目的. 就目前研究的方向和取得的成果来看，现代归纳逻辑是一种哲学逻辑，而概率逻辑是演绎逻辑在数学的某些分支，如概率论、测度论方面的运用，因此是一种应用逻辑. 但是在本世纪的前 50 年中，它们往往是很难区别的. 例如，现在人们公认 J.M. Keynes 是第一个建立概率逻辑公理化系统的人，事实

上，这个系统也是第一个现代归纳逻辑系统．Carnap 也认为概率逻辑就是归纳逻辑．的确，这两种逻辑有许多共同之处．例如，它们都把概率作为基本概念，都把不确定的概然推理作为研究对象，等等．因此许多人都把现代归纳逻辑和概率逻辑归于一大类．我们在此书中将采取灵活的姿态，在能区别的地方将分开研究和讨论，在不能分开的地方，我们就统一进行研究和讨论．

此书的撰写和章节的安排是根据理论和历史相统一的原则．我们将分章介绍和讨论现代归纳逻辑和概率逻辑的一些主要理论． 本书是写给对逻辑感兴趣的读者看的(假设读者有一定的数理逻辑基础知识)，所以在每章，除了介绍建立这样那样的逻辑的目的外，主要致力于逻辑本身的问题，而不过多讨论哲学问题或数学问题．

此书曾作为北京大学哲学系逻辑专业 87 级本科生的教材，由于作者水平有限，错误和不当之处在所难免，希望读者批评指正．

在撰写此书的过程中，得到许多专家和同行，特别是我导师宋文坚教授的指导和帮助，在此向他们表示衷心的感谢．

李小五

1991年7月于北京

目　录

第一章　Keynes 的概率逻辑

J. M. Keynes 是第一个建立概率逻辑公理化系统的人. 他的主要贡献在于：1. 对概率概念作了深入的讨论. 他一方面把概率看作是命题集之间的逻辑关系，另一方面又把它看作是人们对概然命题的合理相信度，这两个思想对后人（如 Carnap）的影响都非常大，但他的缺点是没有提出求初始概率的方法. 2. 建立了第一个概率演算. 他的系统非常丰富，远远超过经典概率论的范围和成果，然而其缺点在于形式化系统不够严谨. 3. 对传统归纳法和类比法进行了大量深入的讨论. 本章着重介绍前两部分的内容.

§1　概率概念的本性

Keynes 认为,我们的知识部分是直接得到的，部分是通过论证得到的. 概率逻辑（logic of probability）就是我们通过论证得到的那部分知识的逻辑. 论证可分两大类，一类称为确定（certain）论证. 例如三段论推理，理想空间的几何等. 还有一类称为合理（rational）论证. 例如,在社会科学和行为科学中,我们经常习惯把我们的合理相信（rational belief，若一个人由于某个荒谬的理由或根本没有任何理由相信一命题,则我们不能称他相信该命题是合理的,虽然该命题可能是真的.）建立在这样的论证上,它们可以在某种程度上是不确定的.

概然(不确定)和确定这两个术语描述了人们对命题的不

同程度的合理相信，这些相信是不同的知识给我们的。任何命题本身非真即假，但是我们关于它的知识却依赖我们的认识环境，因此概然或确定不是命题本身的特性。所以在这个意义上，刻划命题概然程度或确定的概率可以称为主观的.但另一方面，在对逻辑非常重要的意义上，概率不是主观的。一旦给定能确定我们知识的事实，那么概然或确定的含义也就随之客观地固定下来，并且独立于我们的意见。概率逻辑就是研究在给定条件下人们具有的合理相信度，而不是研究人们关于特定具体的事物的实际相信度（Carnap 晚年继承了这一思想）.我们选择什么特定命题集作为我们的论证前提当然是依赖我们特定的主观因素。但是一旦给定这些前提，根据概率逻辑的有效论证能够推出什么是合理相信度（概然的或确定的），这个过程纯粹是客观的。

因此 Keynes 给概率下了如下定义：令前提由任一命题集 h 组成，结论由任意命题集 a 组成。若 h 的知识以程度 r 证明合理相信 a 是正当的（justify），则我们说在 a 和 h 之间存在程度 r 的概率关系（参见[1]，p.4）。用符号表示就是 $a/h=r$。Keynes 认为概率逻辑就是研究这种关系的逻辑。

Keynes 的概率主要有下列两层含义：（1）概率关系是最基本的关系，这种关系不能用更简单的概念来分析或定义，所以 Keynes 的逻辑没有完备的语义解释。（2）有时 Keynes 把概率看作是根据证据 h 对 a 进行论证的可靠程度，这样的概率也称为一个论证的概率。他认为他的理论就是研究从前提到结论的合理（通常又是不确定）论证的一般理论。确定性论证就是从前提可演绎证明结论，此关系就是极大概率关系。不可能性就是从前提不能演绎证明结论，此关系就是极小概率关系，不同的中间值概率分别代表可靠程度不同的论证.这种把概率看作是论证的可靠程度的思想被 L.J. Cohen 继

承.

Keynes认为有些概率是不可测度的，甚至是不可排序的。他说：谁能准确地计算出我们今天步行出门活着回来的概率？有些概率甚至连互相比较大小都不可能。考虑为建立某一概括命题所进行的三组实验。第一组实验次数较多。第二组实验条件变换较多，第三组涉及的概括范围更广。该命题(结论)相对哪一组实验(前提)的概率更大呢？对此显然是无法加以回答的。因此，对概率的测度，只是在一定的条件下，对某些概率才是可能的；概率之间的比较大小，也只是对某些概率才有意义。概率逻辑把概率看作是两命题或命题集之间的逻辑关系，这反映了“概率”最广泛的涵义。既然要建立这种更具一般性的概率逻辑，就只好接受部分概率可测度可比较的思想。

根据上述思想，Keynes提出概率的序理论，认为一概率若自身不是确定的或不可能的，则位于确定性和不可能性之间。同样，若第一个概率和确定性之间还存在第二个概率，则我们就说第二个概率大于第一个概率，这样继续下去就形成了一个有序的概率序列。显然不是所有的概率都能置于同一个序列之中，所以存在多个不同的概率序列，只有那些属于同一序列的概率才能比较。当然在同一序列的每一对概率之间不一定存在另一个概率。总的来说，它依赖无差别原则（Principle of indifference）的运用。该原则是概率逻辑中最重要的规则，它的传统表述是：对某一条件下的某些随机事件，若没有足够的理由认为其中的某些事件比另一些更有可能发生，则应该认为它们具有相等的概率。Keynes认为运用无差别原则必须有一定的限制条件，否则就可能出错。他把这样的限制条件称为“对称性条件”，这个条件也可看作是概率的可测性条件(请参见[1]，第4章)。

§2 概率演算

Keynes 的系统由 19 条初始定义和 7 条初始公理组成，我们在此讨论其中我们感兴趣的定义和公理。这里设 a，b，c，…表示命题，P, Q, R，…表示概率关系，$\wedge$（合取），$\vee$（析取），$\neg$（否定），$\rightarrow$（蕴涵），$\longleftrightarrow$（等价）表示命题联结词.

定义 1.1 在命题 a 和前提 h 之间存在概率关系 $P \Longleftrightarrow a/h = P$.

定义 1.2 P 是确定性关系 $\Longleftrightarrow P = 1$.

定义 1.3 P 是不可能关系 $\Longleftrightarrow P = 0$.

定义 1.4 P 是概率关系但不是确定性关系 $\Longleftrightarrow P<1$.

定义 1.5 P 是概率关系但不是不可能关系 $\Longleftrightarrow P > 0$.

定义 1.6 $a/h = 0 \Longleftrightarrow a \wedge h$ 不一致.

若 a，h 是命题集，则在它们之间使用命题联结词 $\wedge$，$\vee$，$\rightarrow$，$\neg$，$\longleftrightarrow$ 是有问题的. 所以 Keynes 在下列公理 1.1 中做了进一步的限制.

定义 1.7 $b/(h \wedge a) = 1$ 且 $a/(h \wedge b) = 1 \Longleftrightarrow (a \longleftrightarrow b)/h = 1$.

定义 1.8 $(a \wedge b)/h + (a \wedge \neg b)/h = a/h$.

定义 1.9 $(a \wedge b)/h = a/h \cdot b/(h \wedge a) = b/h \cdot a/(h \wedge b)$.

定义 1.10 $P \cdot Q = R$ 且 $Q \neq 0 \Longleftrightarrow P = R/Q$.

定义 1.11 $P + Q = R \Longleftrightarrow P = R - Q$.

定义 1.12 $a_1/(h \wedge a_2) = a_1/h$ 且 $a_2/(h \wedge a_1) = a_2/h \Longleftrightarrow a_1/h$ 和 a_2/h 互相独立.

定义 1.13 $a_1/(h \wedge a_2) = a_1/h \Longleftrightarrow a_2$ 与 a_1/h 不相干.

公理 1.1 若 a 和 h 是命题,或命题的合取，或命题的析取，且 h 不是不一致的合取，则在作为结论的 a 和作为前提的 h 之间,存在唯一的概率关系 P(因此,当我们写作 a/h 时，已预设 h 不是不一致的,即 $h/h\neq 0$，下同)。

公理 1.2 (1)若 $(a\longleftrightarrow b)/h=1$，且 x 是任意命题,则 $x/(h\wedge a)=x/(h\wedge b)$。(2) $x/(a\wedge a)=x/a$。

公理 1.3 (1) $T/h=1$(T 是古典命题演算的重言式)。(2) 若 $a/h=1$,则对任一命题 x,$x/(h\wedge a)=x/h$。

公理 1.4 概率蕴涵关系之间的运算满足普通的算术规则(即概率关系 $P,Q,R,\cdots$表示区间[0,1]之间的实数)。

从上述可见，Keynes 的系统在形式化方面不够严谨.例如他的某些定义(例如定义 1.9)后人常用作公理，他的公理 1.1,后人(如 H. Reichenbach) 常用作限制条件或规则。

下面我们证明和讨论我们感兴趣的定理。

定理 1.1 若 $a\longleftrightarrow b$ 是重言式,则对任一命题 x,$x/a=x/b$。

证明：据公理 1.3,$a\longleftrightarrow b/a=1$ 且 $a\longleftrightarrow b/b=1$，据公理 1.2,

$$x/(a\wedge a)=x/(a\wedge b) \text{ 且 } x/(b\wedge a)=x/(b\wedge b),$$

所以 $x/(a\wedge a)=x/(b\wedge b)$，再据公理 1.2(2)即证。

定理 1.2 若 $(a\longleftrightarrow b)/h=1$，则 $a/h=b/h$。

证明据定义 1.9 和定义 1.7。

定理 1.3 若 a 不是不一致,则 $a/a=1$。

证明：据公理 1.3 和定理1.2,我们有$(a\wedge a)/a=a/a$,据定义 1.9,$a/a\cdot a/(a\wedge a)=a/a$。 据假设，$a/a\neq 0$，所以 $a/(a\wedge a)=1$，再据公理 1.2(2)。

定理 1.4 (否定原则) $a/h+\neg a/h=1$。

证明：据定义 1.8，$(h\wedge a)/h+(h\wedge\neg a)/h=h/h$，再

据定义 1.9，我们有

$$h/h \cdot a/(h \wedge h) + h/h \cdot \neg a/(h \wedge h) = h/h.$$

据公理 1.1，h 不是不一致的，所以据定理 1.3 和公理 1.2(2)，我们有 $a/h + \neg a/h = 1$.

定理 1.5 若 $a/h = 0$，则 $(a \wedge b)/h = 0$；若 $a/h = 0$ 且 $h \wedge b$ 不是不一致的，则 $a/(h \wedge b) = 0$.

证明：据定义 1.9 和假设，$(a \wedge b)/h = 0$；再据定义 1.6，$a/(h \wedge b) = 0$.

Keynes 认为，若一论证不可能，则增加结论或一致地增加前提都不会影响该论证.

定理 1.6 若 $a/h = 1$ 且 $h \wedge b$ 不是不一致的，则 $a/(h \wedge b) = 1$.

证明：据定理 1.4 和假设，$\neg a/h = 0$，据定理 1.5，$\neg a/(h \wedge b) = 0$，再据定理 1.4.

定理 1.7 $(a \vee b)/h = a/h + b/h - (a \wedge b)/h$.

证明：由定义 1.8

$$((a \vee b) \wedge (\neg a \wedge b))/h + ((a \vee b) \wedge \neg(\neg a \wedge b))/h = (a \vee b)/h.$$

据公理 1.3，$((a \vee b) \wedge (\neg a \wedge b) \longleftrightarrow \neg a \wedge b)/h = 1$，

$$((a \vee b) \wedge \neg(\neg a \wedge b) \longleftrightarrow a)/h = 1.$$

从而根据定理 1.2，可得 $(\neg a \wedge b)/h + a/h = (a \vee b)/h$.

据定义 1.9，$b/h \cdot \neg a/(h \wedge b) + a/h = (a \vee b)/h$.

利用定理 1.4 可得

$$b/h - b/h \cdot a/(h \wedge b) + a/h = (a \vee b)/h.$$

由定义 1.9 得，$(a \vee b)/h = a/h + b/h - (a \wedge b)/h$.
定理得证.

定理 1.7 称为普遍析取规则，因为它不要求 $(a \wedge b)/h = 0$（即 a 和 b 相对 h 不相容）.

定理 1.8 若 $(a \wedge b)/h = 0$，则

$$(a \vee b)/h = a/h + b/h.$$

定理 1.9

$$(a_1 \vee a_2 \vee \cdots \vee a_n)/h = \sum_{r=1}^{n}(a_r/h) - \sum_{1 \leqslant r < s \leqslant n}((a_r \wedge a_s)/h)$$

$$+ \sum_{1 \leqslant r < s < t \leqslant n}((a_r \wedge a_s \wedge a_t)/h) + \cdots$$

$$+ (-1)^{n-1}(a_1 \wedge a_2 \wedge \cdots \wedge a_n)/h.$$

证明：据定理 1.7，且对 n 运用归纳法。

定理 1.10 若对任意 $s,t \leqslant n$，都有 $(a_s \wedge a_t)/h=0$（即 $a_1,\cdots,a_n$ 相对 h 互斥），则

$$(a_1 \vee \cdots \vee a_n)/h = \sum_{r=1}^{n}(a_r/h).$$

定理 1.11 若对任意 $s,t \leqslant n$，都有 $(a_s \wedge a_t)/h = 0$ 且 $(a_1 \vee \cdots \vee a_n)/h = 1$（即 $a_1, \cdots, a_n$ 相对 h 是一互斥且联合穷举的选择集），则

$$\sum_{r=1}^{n}(a_r/h) = 1.$$

定理 1.12 若假设条件与定理 1.11 相同，则

$$a/h = \sum_{r=1}^{n}((a_r \wedge a)/h).$$

证明：若 $h \wedge a$ 不一致，据定义 1.6，则 $a/h = 0$（因为根据预设 h 一致）。因此，据定理 1.5，立得

$$\sum_{r=1}^{n}((a_r \wedge a)/h) = 0.$$

所以下面假设 $h \wedge a$ 不是不一致的. 据题设和定理 1.6 可知，$(a_1 \vee \cdots \vee a_n)/(h \wedge a)=1$. 据题设和定理 1.5 可知，$(a_s \wedge a_t)/$

$(h\wedge a)=0$。因此再据定理 1.10 和定理 1.11 就有

$$\sum_{r=1}^{n}(a_r/(h\wedge a))=(a_1\vee\cdots\vee a_n)/(h\wedge a)=1,$$

由定义 1.9,就有

$$\sum_{r=1}^{n}((a_r\wedge a)/h)=a/h\cdot\sum_{r=1}^{n}(a_r/(h\wedge a))=a/h.$$

定理 1.13 $(h\to a)/h=a/h$.

证明：据定理 1.7,有

$$(h\to a)/h=\neg h/h+a/h-(\neg h\wedge a)/h,$$

而据定理 1.3 和预设，有 $\neg h/h=0$，因此据定理 1.5，有 $(\neg h\wedge a)/h=0$。所以 $(h\to a)/h=a/h$.

定理 1.14 $(a\vee b)/h=0\Longleftrightarrow a/h=0$ 且 $b/h=0$. ("$\Longleftrightarrow$"表示"当且仅当",下同). Keynes 本人对此定理的证明似乎有问题,我们重新证明,为此先证一重要的引理。

引理 若 $a\to b$ 是重言式,则 $a/h\leqslant b/h$.

证明:

从重言式可知，$(a\to b)\to(b\longleftrightarrow(b\wedge\neg a)\vee a)$.

因此据假设和分离规则，可知，$(b\longleftrightarrow(b\wedge\neg a)\vee a)$.

据公理 1.3 和定理 1.2，可得 $b/h=((b\wedge\neg a)\vee a)/h$.

据公理 1.3,得 $\neg((b\wedge\neg a)\wedge a)/h=1$.

从而据定理 1.4，可得 $((b\wedge\neg a)\wedge a)/h=0$.

应用定理1.8,则有$((b\wedge\neg a)\vee a)/h=(b\wedge\neg a)/h+a/h$,因此，$b/h-a/h=(b\wedge\neg a)/h$.

据定义 1.3 和定义 1.5，可得 $0\leqslant(b\wedge\neg a)/h$.

联立上两式得

$$0\leqslant b/h-a/h,$$

因此引理得证。

现在我们来证定理 1.14，据上述引理,我们有

$$(a \wedge b)/h \leqslant (a \vee b)/h,$$

据假设 $(a \vee b)/h = 0$，有 $(a \wedge b)/h = 0$。再据定理 1.7，$a/h + b/h = 0$，因此 $a/h = 0$ 且 $b/h = 0$。另一方面，若 $a/h = 0$ 且 $b/h = 0$，据定理 1.5，$(a \wedge b)/h = 0$，再据定理 1.7，$(a \vee b)/h = 0$。

定理 1.15 若 $a/h = 1$，则 $(b \to a)/h = 1$。

证明：据假设、定理 1.4 和 1.5，易证 $(\neg b \wedge \neg a)/h = 0$，因此

$$\begin{aligned}
&(b \to a)/h \\
&\quad = (\neg b \vee a)/h \quad \text{(据公理 1.3 和定理 1.2)} \\
&\quad = \neg b/h + a/h - (\neg b \wedge a)/h \quad \text{(据定理 1.7)} \\
&\quad = a/h + (\neg a \wedge \neg b)/h \quad \text{(据定义 1.8)} \\
&\quad = 1.
\end{aligned}$$

定理 1.15 说明概率演算保留类似蕴涵悖论的结果：一确定命题为任一命题所蕴涵也同样是确定的。

定理 1.16 若 $a/(h_1 \wedge h_2) = a/h_1$ 且 $\neg h_2/h_1 \neq 0$，则 $a/(h_1 \wedge \neg h_2) = a/h_1$。

证明：据定义 1.8，1.9 和假设，有

$$\begin{aligned}
a/h_1 &= (a \wedge h_2)/h_1 + (a \wedge \neg h_2)/h_1 \\
&= h_2/h_1 \cdot a/(h_1 \wedge h_2) + \neg h_2/h_1 \cdot a/(h_1 \wedge \neg h_2) \\
&= h_2/h_1 \cdot a/h_1 + \neg h_2/h_1 \cdot a/(h_1 \wedge \neg h_2).
\end{aligned}$$

因此，有 $a/h_1(1 - h_2/h_1) = \neg h_2/h_1 \cdot a/h_1 \wedge \neg h_2$，因为

$$\neg h_2/h_1 \neq 0,$$

所以据定理 1.4 和公理 1.4，我们有 $a/h_1 = a/(h_1 \wedge \neg h_2)$。

定理 1.16 说明，若一命题与一论证不相干，则它的矛盾命题也与该论证不相干。

定理 1.17 若 $a/(h \wedge h_1) > a/h$ 且 $h_1 \wedge h$ 不是不一致的，则 $h_1/(h \wedge a) > h_1/h$。

证明：据定义 1.9 和假设，

$$a/h \cdot h_1/(h \wedge a) = h_1/h \cdot a/(h \wedge h_1) > h_1/h \cdot a/h.$$

假设 $a \wedge h$ 不是不一致的，即 $a/h \neq 0$，则有 $h_1/(h \wedge a) > h_1/h$. 若 $a \wedge h$ 不一致，则 $a/h = 0$. 因为据题设，$h_1 \wedge h$ 不是不一致的，所以据定理 1.5，$a/(h_1 \wedge h) = 0$，因此 $a/h = a/(h_1 \wedge h)$，与题设矛盾。

定理 1.17 说明，若 h_1 对论证 a/h 是有利相干（即 $a/(h \wedge h_1) > a/h$)，则 a 对论证 h_1/h 也如此。Keynes 认为定理 1.17 构成一般为人们接受的一种形式证明原则：若假说 h_1 有助于解释现象 a，则 a 的事实支持 h_1 的实现。

定理 1.18 若 a_1/h 和 a_2/h 互相独立，则 $(a_1 \wedge a_2)/h = a_1/h \cdot a_2/h$（据定义 1.9 和 1.12）。

定理 1.19 (Bayes 定理) 若 $a_1, \cdots, a_n$ 互斥且联合穷举且 $b/h > 0$，则

$$a_m/(h \wedge b) = \frac{a_m/h \cdot b/(h \wedge a_m)}{\sum_{r=1}^{n}(a_r/h \cdot b/(h \wedge a_r))},$$

其中 $m = 1, 2, \cdots, n$.

证明：据定义 1.9，

$$b/h \cdot a_m/(h \wedge b) = a_m/h \cdot b/(h \wedge a_m).$$

因此，据假设和定理 1.12 以及定义 1.9，就能得到本定理。

Bayes 定理在归纳逻辑中很重要。Keynes 认为，可以用自然语言来说明此定理。令 b 表示事件 B 的发生，$a_1, \cdots, a_n$ 是相对背景材料 h 互斥且联合穷举的假说集，分别表示 B 的 n 个可能的原因 $A_1, \cdots, A_n$ 的存在。于是，当我们不知事件 B 是否发生时，$a_1/h, \cdots, a_n/h$ 分别是 $A_1, \cdots, A_n$ 存在的先验概率，$b/(h \wedge a_1), \cdots, b/(h \wedge a_n)$ 分别表示假说 $a_1, \cdots,$

a_n 对结果 b 的预测度，则

$$\frac{a_1/h \cdot b/(h \wedge a_1)}{\sum_{r=1}^{n}(a_r/h \cdot b/(h \wedge a_r))}, \cdots, \frac{a_n/h \cdot b/(h \wedge a_n)}{\sum_{r=1}^{n}(a_r/h \cdot b/(h \wedge a_r))}$$

分别表示已知事件 B 发生后，$A_1, \cdots, A_n$ 存在的概率. 这里的初始条件(即 $h \wedge b$ 不是不一致的)只是为了保证事件 B 的发生根据初始材料 h 至少是可能的. Keynes 把此定理称为逆概定理. 他认为概率演算可以用在两类问题上. 一类是由因到果，即给定原因，我们演绎出结果来；另一类则相反，给定结果，我们探索原因，逆概定理就是用在后一类问题上的.

§3 对归纳法的讨论

Keynes 认为概率逻辑是研究论证的逻辑. 这种论证的最重要的类型是依赖归纳法和类比法的论证. 几乎所有的经验科学都依赖这样的论证. 在经验论证中，存在两类归纳概括，一种称为全称归纳法 (universal induction)，枚举归纳法是全称归纳法的一种. 这种归纳法断定的结论要求是全称的，若发现例外，结论就推翻. 但在大多数场合，我们喜欢另一种归纳法，它能逐渐接近一般我们能够依赖的规律，但是无论怎样恰当地建立，这种归纳法只能得到概然的规律. 这类归纳法可以称为归纳关联 (inductive correlation)，即统计归纳法. 例如，我们根据已观察到各种各样的天鹅都是白的，推出结论“所有的天鹅都是白的”，那么我们运用的是全称归纳法. 若我们根据已观察到的这些天鹅是白的且那只天鹅是黑的，推出结论绝大多数天鹅是白的，或一只天鹅是白的的概率是多少，那么我们用的是归纳关联.

Keynes 认为应该强调归纳法和概率逻辑之间的基本联

系，许多人已认识到，根据归纳论证达到的结果是概然的。但很少有人清楚地认识到，在严格意义上，每种归纳法的有效性，不是依赖事物，而是依赖概率关系的存在。归纳论证不是确证某事物如此等等，而是相对某证据确证存在一支持它的概率。

Keynes 认为在两类归纳法中，全称归纳法描述了较简单或更基本的问题。由于 Keynes 对归纳法的讨论主要是哲学的，所以我们不在此展开，只举一个我们感兴趣的逻辑问题，即简单枚举法如何确定全称概括句的概率。

令 h 表示在科学研究中的背景材料，g 表示我们要建立的全称概括句，$x_1, \cdots, x_n$ 是 g 的特例，因此我们有

$$x_1/(h \wedge g) = \cdots = x_n/(h \wedge g) = 1.$$

现在的问题是确定 $g/(h \wedge x_1 \wedge \cdots \wedge x_n)$ 的概率，即给定背景材料 h 和 n 个 g 的特例，求 g 的概率。假设不存在先验的理由来期望任一特例的出现比另一特例的出现更令人信服，也即假设相对 h 的无差别原则

$$x_1/h = \cdots = x_n/h.$$

令 $g/(h \wedge x_1 \wedge \cdots \wedge x_n) = p_n$，$x_{n+1}/(h \wedge x_1 \wedge \cdots \wedge x_n) = y_{n+1}$，假设所有的 $y_{n+1} \neq 0$，于是，根据假设和定义 1.9，定理 1.6 以及 $y_n > 0$，我们有

$$p_n/p_{n-1} = \frac{g/(h \wedge x_1 \wedge \cdots \wedge x_n)}{g/(h \wedge x_1 \wedge \cdots \wedge x_{n-1})}$$

$$= \frac{(g/(h \wedge x_1 \wedge \cdots \wedge x_n))(x_n/(h \wedge x_1 \wedge \cdots \wedge x_{n-1})}{(g/(h \wedge x_1 \wedge \cdots \wedge x_{n-1}))(x_n/(h \wedge x_1 \wedge \cdots \wedge x_{n-1})}$$

$$= \frac{(g/(h \wedge x_1 \wedge \cdots \wedge x_{n-1}))(x_n/(h \wedge x_1 \wedge \cdots \wedge x_{n-1} \wedge g))}{(g/(h \wedge x_1 \wedge \cdots \wedge x_{n-1}))(x_n/(h \wedge x_1 \wedge \cdots \wedge x_{n-1}))}$$

$$= \frac{x_n/(h \wedge x_1 \wedge \cdots \wedge x_{n-1} \wedge g)}{x_n/(h \wedge x_1 \wedge \cdots \wedge x_{n-1})} = 1/y_n$$

所以，$p_n = p_{n-1}/y_n = p_0/(y_1y_2\cdots y_n)$（其中 $p_0 = g/h$ 称为 g 的先验概率）.因此，若 $y_n < 1$，即 x_n 和 $h\wedge x_1\wedge\cdots\wedge x_{n-1}$ 没有推出关系，则 $p_n > p_{n-1}$. 据定义 1.9 和假设，我们还可以进一步考虑：

$$
\begin{aligned}
&(x_1\wedge\cdots\wedge x_n)/h \\
&\quad = ((x_1\wedge\cdots\wedge x_{n-1})/h)(x_n/(h\wedge x_1\wedge\cdots\wedge x_{n-1})) \\
&\quad = ((x_1\wedge\cdots\wedge x_{n-1})/h)\cdot y_n \\
&\quad\quad \vdots \\
&\quad = y_ny_{n-1}\cdots y_1.
\end{aligned}
$$

因此，根据定义 1.8，1.9，定理 1.4，1.6，我们有

$$
\begin{aligned}
p_n &= p_0/y_ny_{n-1}\cdots y_1 \\
&= p_0/((x_1\wedge\cdots\wedge x_n)/h) \\
&= p_0/(((x_1\wedge\cdots\wedge x_n\wedge g)/h) \\
&\qquad + ((x_1\wedge\cdots\wedge x_n\wedge\neg g)/h)) \\
&= p_0/((g/h)((x_1\wedge\cdots\wedge x_n)/(h\wedge g)) \\
&\qquad + ((x_1\wedge\cdots\wedge x_n\wedge\neg g)/h)) \\
&= p_0/(g/h + (x_1\wedge\cdots\wedge x_n\wedge\neg g)/h) \\
&= p_0/(g/h + \neg g/h\cdot(x_1\wedge\cdots\wedge x_n)/(h\wedge\neg g)) \\
&= p_0/(p_0 + (1-p_0)((x_1\wedge\cdots\wedge x_n)/(h\wedge\neg g)))
\end{aligned}
$$

当 n 增大时，若 $(1-p_0)((x_1\wedge\cdots\wedge x_n)/(h\wedge\neg g))\to 0$，则 $p_n\to 1$. 因此 Keynes 认为若每一特例都能从 g 中必然推出，则每一增加的特例都能提高 g 的概率，只要这新的特例不能从前面的特例中确定性地预见到，即 $0<x_n/(h\wedge x_1\wedge\cdots\wedge x_{n-1})<1$，因此可以证明，当特例的数目增加使得 $(1-p_0)((x_1\wedge\cdots\wedge x_n)/(h\wedge\neg g))\to 0$，则归纳论证的概率趋于确定性作为它的极限，这恰恰是简单枚举法的推理过程。

关于类比法和统计归纳法，请参见[1]。

§4 Keynes 对我们的启发

Keynes 在§2给出的公理化系统实际上是一形式语义学系统，并没有建立相应的语法系统（后人，包括 Carnap，也有类似问题），这在搞逻辑的人看来自然是不完备的．在此我们根据 Keynes 的思想，利用现有的成果（主要是可能世界语义学和测度论的成果），试图建立一个公理化的语法系统．希望这种尝试能引起读者的兴趣．

我们考虑的语言是

初始命题集 $PR=\{p_i:i=1,2,\cdots\}$,

命题联结词：$\neg$, $\wedge$,

合理相信算子：$\boldsymbol{B}_i^r$，其中 $j\in J$, $r\in[0,1]$.

其中，J 是一认知主体集，若 φ 是公式，则 $\boldsymbol{B}_j^r\varphi$ 表示认知主体 j 对命题 φ 的合理相信度不小于实数 $r\in[0,1]$．为了简单起见，我们只考虑 $J=\{j\}$ 的情况，所以下面我们省略合理相信算子 $\boldsymbol{B}$ 的下标．

（合式）公式的形成规则除了通常的，还增加下列规则：若 φ 是公式，则 $\boldsymbol{B}^r\varphi$ 也是公式 $(r\in[0,1])$．

除了通常对 $\vee$, $\rightarrow$, $\longleftrightarrow$ 的定义，还引入

$$\boldsymbol{B}^{[r]}\varphi\triangleq\boldsymbol{B}^r\varphi\wedge\boldsymbol{B}^{1-r}(\neg\varphi),$$

$\boldsymbol{B}^{[r]}\varphi$ 可以直观地理解为对 φ 的合理相信度 $=r$.

概率模型 $\mathfrak{U}$ 是下列四元组的结构

$$\mathfrak{U}=\langle W,R,I,m^w\rangle_{w\in W},$$

其中

（1）$W\neq\varnothing$（W 可以理解为可能世界集）．

（2）$R\subseteq W\times W$（R 可以理解为认知通达关系）．

（3）$I:PR\rightarrow\mathscr{P}(W)$，其中 $\mathscr{P}(W)$ 是 W 的幂集，I 是

二值真值赋值，因为根据 Keynes，任何命题本身非真即假。

(4) 对任一 $w \in W$，先令 $AC(w) = \{w' \in W : \langle w, w' \rangle \in R\}$，我们称 $AC(w)$ 是从 w 通达别的可能世界的认知通达集，则 m^w 是 $AC(w)$ 上的概率测度。具体说 $\langle AC(w), \mathscr{F}(AC(w)), m^w \rangle$ 是一概率空间，其中 $\mathscr{F}(AC(w))$ 是 $AC(w)$ 的子集作成的 σ 代数（即 $AC(w) \in \mathscr{F}(AC(w))$ 且 $\mathscr{F}(AC(w))$ 在补和可数并下封闭），$\mathscr{F}(AC(w))$ 的元素称为可测的。m^w 是测度函数 $m^w : \mathscr{F}(AC(w)) \to [0,1]$ 使得下列条件成立：

① 非负性和正规性：对任意 $A \in \mathscr{F}(AC(w))$，$m^w(A) \geq 0$；若 $AC(w) \neq \varnothing$，则 $m^w(AC(w)) = 1$，否则

$$m^w(AC(w)) = 0.$$

② 单调性：对任意 A，$B \in \mathscr{F}(AC(w))$，若 $A \subseteq B$，则 $m^w(A) \leqslant m^w(B)$。

③ σ 可加性：对任意 $A_i \in \mathscr{F}(AC(w))$（其中 $i = 1, 2, \cdots$）使得当 $i \neq j$ 时，$A_i \cap A_j = \varnothing$，则

$$m^w\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} m^w(A_i).$$

令 φ 是任意公式，我们用 $\mathfrak{U}, w \models \varphi$ 表示 φ 在模型 $\mathfrak{U}$ 的论域 W 的元素 w 上成立（真）。令

$$TR(w, \varphi) = \{w' \in AC(w) : \mathfrak{U}, w' \models \varphi\},$$

即 $TR(w, \varphi)$ 是 $AC(w)$ 中使 φ 成立的可能世界集。因此满足关系可以定义如下：

(1) $\mathfrak{U}, w \models p_i \Longleftrightarrow w \in I(p_i)$.

(2) $\mathfrak{A}$，$w \models \neg\varphi \Longleftrightarrow \mathfrak{A}$，$w \not\models \varphi$（即 φ 在 $\mathfrak{U}$ 的 w 上不成立）。

(3) $\mathfrak{U}, w \models \varphi \wedge \phi \Longleftrightarrow \mathfrak{U}$，$w \models \varphi$ 且 $\mathfrak{U}, w \models \phi$.

(4) $\mathfrak{U}, w \models B^r(\varphi) \Longleftrightarrow (m^w)^*(TR(w, \varphi)) \geqslant r$.

其中内测度 $(m^w)^*$ 由 m^w 导出，即对任意 $A\subseteq AC(w)$,

$$(m^w)^*(A)=\sup\{m^w(B): B\subseteq A 且 B\in\mathscr{F}(AC(w))\}.$$

若A不可测 ($A\notin\mathscr{F}(AC(w))$)，则$A$的内测度是被$A$所包含的最大可测集的测度。

我们采用内测度是避免 $TR(w,\varphi)$ 不可测。显然这样定义的 m^* 是m的扩张。事实上，它在W的每一子集上都有定义。另外还需要下列形成规则

若 φ,ϕ 是公式，则 $B^r(\varphi,\phi)$ 也是公式。

我们还可以有下列满足关系

(5) $\mathfrak{U},w\models B^r(\varphi,\phi)\Longleftrightarrow(m^w)^*(TR(w,\varphi\wedge\phi))/((m^w)^*\cdot(TR(w,\phi)))\geqslant r$，若 $(m^w)^*(TR(w,\phi))\neq 0$.

直观地说，$B^r(\varphi,\phi)$表示我们根据已知的知识ϕ，对φ的合理相信度$\geqslant r$。这样表示的二元合理相信算子更接近 Keynes 的原意，用 Keynes 的符号表示，$B^r(\varphi,\phi)$ 就是 $\varphi/\phi\geqslant r$.

相应地我们定义

$$B^{[r]}(\varphi,\phi)\triangleq B^{[s]}(\varphi\wedge\phi)\wedge B^{[t]}(\phi),$$

其中 $t>0$ 且 $s/t=r$.

现在我们可以给出概括Keynes系统的公理化语法系统，其中实数 $r,s\in[0,1]$.

公理 1.5 古典命题演算的重言式及其代入特例。

公理 1.6 $B^0\varphi$.

公理 1.7 (单调性) $B^r\varphi\to B^s\varphi$，其中 $r\geqslant s$.

公理 1.8 $B^1(\varphi\to\phi)\to(B^1\varphi\to B^1\phi)$.

公理 1.9 (乘法公理)

(1) $B^{[r]}\varphi\wedge B^{[s]}(\phi,\varphi)\to B^{[r\cdot s]}(\varphi\wedge\phi)$.

(2) $B^{[r]}(\varphi\wedge\phi)\wedge B^{[s]}\varphi\to B^{[r/s]}(\phi,\varphi)$，其中 $s>0$.

(3) $B^{[r]}(\varphi\wedge\phi)\wedge B^{[s]}(\varphi,\phi)\to B^{[r/s]}\phi$，其中 $s>0$.

规则 1.1 $\varphi\to\phi,\varphi/\phi$.

规则 1.2 $\varphi / \boldsymbol{B}^{1}\varphi$.

规则 1.3 $\varphi \to \psi / \boldsymbol{B}^{r}\varphi \to \boldsymbol{B}^{r}\psi$.

规则 1.4 （加法规则）

（1） $\neg(\varphi \wedge \psi) / \boldsymbol{B}^{[r]}\varphi \wedge \boldsymbol{B}^{[s]}\psi \to \boldsymbol{B}^{[r+s]}(\varphi \vee \psi)$.

（2） $\neg(\varphi \wedge \psi) / \boldsymbol{B}^{[r]}\varphi \wedge \boldsymbol{B}^{[s]}(\varphi \vee \psi) \to \boldsymbol{B}^{[s-r]}\psi$.

下面我们来证明几个定理.

定理 1.20 $\boldsymbol{B}^{[r]}\varphi \longleftrightarrow \boldsymbol{B}^{[1-r]}(\neg\varphi)$.

证明:

据公理 1.5，$\neg(\varphi \wedge \neg\varphi)$.

从而根据规则 1.4，$\boldsymbol{B}^{[r]}\varphi \wedge \boldsymbol{B}^{[s]}(\varphi \vee \neg\varphi) \to \boldsymbol{B}^{[s-r]}(\neg\varphi)$.
此外我们利用规则 1.2，可得 $\boldsymbol{B}^{[1]}(\varphi \vee \neg\varphi)$（严格说据规则 1.2 的平凡变形）.根据规则 1.1，我们有 $\boldsymbol{B}^{[r]}\varphi \to \boldsymbol{B}^{[1-r]}(\neg\varphi)$.

注意 $\boldsymbol{B}^{[1]}(\varphi \vee \neg\varphi) \wedge \boldsymbol{B}^{[r]}\varphi \to \boldsymbol{B}^{[1-r]}(\neg\varphi)$ 是规则 1.4（2）的一特例. 因此我们有 $\boldsymbol{B}^{[1-r]}(\neg\varphi) \to \boldsymbol{B}^{[r]}\varphi$. 定理得证.

请读者比较定理 1.20 和 §2 的定理 1.4.

在证明下面的定理之前,我们先证导出规则:

$$\varphi \to \psi / \boldsymbol{B}^{[r]}\varphi \to \boldsymbol{B}^{[r]}\psi.$$

证明: 据规则 1.3，有

$$\varphi \to \psi / \boldsymbol{B}^{r}\varphi \to \boldsymbol{B}^{r}\psi,$$

$$\varphi \to \psi / \boldsymbol{B}^{1-r}(\neg\varphi) \to \boldsymbol{B}^{1-r}(\neg\psi).$$

因此，$\varphi \to \psi / \boldsymbol{B}^{r}\varphi \wedge \boldsymbol{B}^{1-r}(\neg\varphi) \to \boldsymbol{B}^{r}\psi \wedge \boldsymbol{B}^{1-r}(\neg\psi)$.

定理 1.21 $B^{[r]}(\varphi \wedge \psi,\ \theta) \wedge B^{[s]}(\varphi \wedge \neg\psi,\ \theta) \to B^{[r+s]}(\varphi, \theta)$.

证明: 据公理 1.5 我们有，

$$\neg((\varphi \wedge \psi \wedge \theta) \wedge (\varphi \wedge \neg\psi \wedge \theta)).$$

利用加法规则,导出规则和公理 1.5 可得

$$\boldsymbol{B}^{[t\cdot r]}(\varphi \wedge \psi \wedge \theta) \wedge \boldsymbol{B}^{[t\cdot s]}(\varphi \wedge \neg\psi \wedge \theta) \to$$

$$B^{[t(r+s)]}((\varphi\wedge\psi\wedge\theta)\vee(\varphi\wedge\neg\psi\wedge\theta)),$$

$$B^{[t(r+s)]}((\varphi\wedge\psi\wedge\theta)\vee(\varphi\wedge\neg\psi\wedge\theta))\to$$
$$B^{[t(r+s)]}(\varphi\wedge\theta).$$

联立上述两式，并且利用三段论，我们有

$$B^{[t\cdot r]}(\varphi\wedge\psi\wedge\theta)\wedge B^{[t\cdot s]}(\varphi\wedge\neg\psi\wedge\theta)\to$$
$$B^{[t(r+s)]}(\varphi\wedge\theta).$$

再据古典命题演算的附加规则可得

$$B^{[t]}(\theta)\wedge B^{[t\cdot r]}(\varphi\wedge\psi\wedge\theta)\wedge B^{[t]}(\theta)\wedge$$
$$B^{[t\cdot s]}(\varphi\wedge\neg\psi\wedge\theta)\to B^{[t]}(\theta)\wedge$$
$$B^{[t(r+s)]}(\varphi\wedge\theta),\ r>0.$$

因此

$$B^{[r]}(\varphi\wedge\psi,\theta)\wedge B^{[s]}(\varphi\wedge\neg\psi,\theta)\to B^{[r+s]}(\varphi,\theta).$$

定理得证。

请比较定理 1.21 和 § 2 的定义 1.8.

从上述两定理的证明，我们注意到关于一元相信算子的定理和关于二元相信算子的定理的证明途径大致相同，只是后者更麻烦一些。

我们还可以在这个系统中证明以下结论

$$B^r(\varphi\wedge\psi)\to B^r\varphi\wedge B^r\psi,$$

$$B^1(\varphi\wedge\psi)\longleftrightarrow B^1\varphi\wedge B^1\psi,$$

$$B^{[0]}(\varphi,\theta)\to B^{[0]}(\varphi\wedge\psi,\theta),$$

$$B^{[0]}(\varphi,\theta)\wedge B^r(\psi,\theta)\to B^{[0]}(\varphi,\psi\wedge\theta),r>0,$$

$$B^{[0]}(\varphi\vee\psi,\theta)\longleftrightarrow B^{[0]}(\varphi,\theta)\wedge B^{[0]}(\psi,\theta),$$

后面三个定理可以分别与 §2 的定理 1.5 和定理 1.14 作比较。

显然，我们能得到上述系统的可靠性定理，即若 φ 是上述系统的定理，则它在所有的概率模型中真。关于逆定理，即

完全性定理是否成立的问题似乎不易解决，原因在于通常模态逻辑使用的典范模型的方法在此失效。我们希望读者能够关注这一问题，因为猜想起来，完全性定理应该是有的，而且这是一个至关重要的问题。

第二章　Reichenbach 的概率逻辑与归纳逻辑

本世纪 20 年代，R. von Mises 提出了频率主义的概率理论，把概率看作某个随机聚合的子序列的相对频率的极限. H. Reichenbach 继承和发展了这种思想. 他在概率逻辑和归纳逻辑方面的主要贡献有：(1) 把概率蕴涵号引入到带等词的一阶谓词演算中，形成了相当丰富的概率演算，并用频率解释和认定理论作为它们的语义模型，初步建立了既有语法系统又有语义模型的概率演算系统. (2) 阐述了概率逻辑的本性，建立了概率逻辑的真值表. 把概率逻辑看作是取连续值的多值逻辑，分析了这样的多值逻辑与古典二值逻辑的关系. (3) 深入讨论了归纳法的分类和作用，以及归纳逻辑的本性，归纳逻辑、概率逻辑与演绎逻辑的关系.

§1　概 率 演 算

H. Reichenbach 的概率演算是在带等词的狭谓词演算的基础上，增加概率蕴涵号和引人新的公理而建立起来的. 他利用了罗素的系统，采用类演算的符号来表示公理和别的公式. 他的概率演算有三个系统：初等概率演算、概率序列的序理论和高阶概率理论，其中以前者最为基本，所以我们在此只讨论它，其它可参见[2].

1. 概率蕴涵关系的引入

Reichenbach 用 $a,b,c,\cdots$ 表示命题变项，$f,g,h,\cdots$ 表示谓词，$F,G,H,\cdots$ 表示类，$\neg,\vee,\wedge,\rightarrow,\leftrightarrow$ 定义如通

常。Reichenbach 用谓词来定义类

$$x \in F \triangleq f(x),$$
$$x, y \in F \triangleq f(x, y),$$
$$x \in \bar{F} \triangleq \neg(x \in F),$$
$$x \in (F \to G) \triangleq x \in (\bar{F} \vee G).$$

通过这样的定义，谓词演算的普遍有效式就对应类演算的普遍有效式，其中全类 V（即所有东西的类）看作是真命题的外延，空类 $\varnothing$ 看作是假命题的外延。

Reichenbach 认为要在类演算的基础上建立概率演算，首先要把概率陈述句形式化。考虑一个典型的概率陈述句"若 a 真，则 b 以程度 p 概然"。它肯定了命题 a 和 b 之间存在概率蕴涵关系，用符号来表示就是 $a \underset{p}{\ni} b$。概率蕴涵联结的项通常是事件。令 x 表示事件"此骰子掷出"，y 表示事件"此骰子在桌上停下后 1 朝上"，这就意味事件 y 被看作属于某个类 B。我们关心的是某个类，至于该骰子滚到桌子的什么地方，或者它的边朝什么方向，都是无关紧要的，只考虑 1 朝上就行了。这对事件 x 也同样，我们不考虑用什么力来掷它，或传给它是什么角动量，只要求事件 x 就是掷一骰子，所以它也是属于某个类 A 的。因此，我们可以用以下形式来刻划这个概率陈述句

$$x \in A \underset{p}{\ni} y \in B. \tag{2.1}$$

但这个表述还需修改：类的元素必须用某种次序给定，例如时间的次序。换句话说，事件 x 属于可数序列 $\langle x_i : i = 1, 2, \cdots \rangle$（以后简记为 $\langle x_i \rangle$），同时事件 y 属于对应的序列 $\langle y_i : i = 1, 2, \cdots \rangle$（以后简记为 $\langle y_i \rangle$）。在两序列的元素之间存在一一对应的关系，这用相同的下标表示。我们只能断定对应元 x_i 和 y_i 之间的概率蕴涵句，所以（2.1）修改为

$$x_i \in A \underset{p}{\supset} y_i \in B. \tag{2.2}$$

但(2.2)也没有完全表达概率陈述句的形式，我们还必须进一步断定：对每一 x_i, y_i，相应的概率蕴涵句都成立，因此(2.2)用下列全称闭包(2.3)来代替

$$(x_i)(y_i)\left(x_i \in A \underset{p}{\supset} y_i \in B\right). \tag{2.3}$$

表达式(2.3)表示了该概率陈述句的最终形式．(2.3)也说明概率蕴涵可以看作是类之间的一种关系．我们把类 A 称为参考类，类 B 称为属性类．为了方便，(2.3)可以缩写为

$$\left(A \underset{p}{\supset} B\right). \tag{2.4}$$

为了和概率论的记号一致，我们引入以下定义

$$(P(A,B)=p) \triangleq \left(A \underset{p}{\supset} B\right).$$

不过要注意，Reichenbach 的 $P(A, B)$ 恰是现行概率论的 $P(B,A)$.

2. 公理和规则

到目前为止我们还没有给出一种确定任意两个类之间是否存在概率关系的方法，即求初始概率的方法(下一节将讨论这种方法)．所以我们在此只研究如何从给定的概率蕴涵句推出新的概率蕴涵句的问题，先引入以下公设：

存在规则：若根据下列概率演算，概率蕴涵句 $\left(A \underset{p}{\supset} B\right)$ 的数值 p 能被给定的概率句确定，则该概率蕴涵句 $\left(A \underset{p}{\supset} B\right)$ 存在。

Reichenbach 认为存在规则不是概率演算的公理，它只是一个用元语言表述的规则，类似经典逻辑中的分离规则和

代入规则.

Reichenbach 的概率演算的语言是一阶带等词的谓词演算的语言加上概率蕴涵号$\underset{p}{\ni}$，$p\in[0,1]$. 他的公理系统可以看作是刻划概率蕴涵的性质的公式系统.

公理 2.1 （单一性公理）

$$p\neq q\rightarrow\left(\left(A\underset{p}{\ni}B\right)\wedge\left(A\underset{q}{\ni}B\right)\longleftrightarrow(x_i)\neg(x_i\in A)\right).$$

公理 2.2 （正规性公理）

(1) $(A\rightarrow B)\rightarrow(\exists p)\left(\left(A\underset{p}{\ni}B\right)\wedge(p=1)\right).$

(2) $(\exists x_i)(x_i\in A)\wedge\left(A\underset{p}{\ni}B\right)\rightarrow(p\geqslant 0).$

公理 2.3 （加法公理）

$$\left(A\underset{p}{\ni}B\right)\wedge\left(A\underset{q}{\ni}C\right)\wedge(A\rightarrow\neg(B\wedge C))\rightarrow$$
$$(\exists r)\left(\left(A\underset{r}{\ni}B\vee C\right)\wedge(r=p+q)\right).$$

公理 2.4 （乘法公理）

$$\left(A\underset{p}{\ni}B\right)\wedge\left(A\wedge B\underset{q}{\ni}C\right)\rightarrow$$
$$(\exists r)\left(\left(A\underset{r}{\ni}B\wedge C\right)\wedge(r=p\cdot q)\right).$$

公理 2.5 带等词的一阶谓词演算的全部普遍有效式(同构于类演算的全部普遍有效式). Reichenbach 公理系统的规则与带等词的一阶谓词演算的规则相同，只是对概率蕴涵号的前后件不能代入概率蕴涵式.

上述系统也是一种公理化语法系统. 公理 2.1 表述了概率度的单一性，这是用一种归谬法形式来表述的. 公理 2.2 说明概率度 $p\in[0,1]$，这种正规性要求类A非空. 乘法公理可

以与第一章的定义 1.9 相比较.

为了方便，Reichenbach 也用 P 记号来表示公理，但在我们看来这种变换似乎是不严格的. 例如，把加法公理用 P 记号来表示,就有

若 $A \to \neg(B \wedge C)$ 是普遍有效式,则

$$P(A, B \vee C) = P(A, B) + P(A, C).$$

这样的变换不仅扩大了含义，而且把语法公理换成语义公理(因为除非不把 P 看作是一阶语言中的函数符号,否则上述等式已不是一阶公式).

从上述系统可以推导出许多关于概率的定理来，请读者参考第一章自证.

§2 频率解释模型

Reichenbach 的频率解释模型是把概率解释为无穷序列的相对频率的极限. 不失一般性,我们并不假设序列 $\langle x_i \rangle$ 的所有元都属于 A. 例如掷一硬币的序列可能相间另一硬币的情况. 同样,在对应的序列 $\langle y_i \rangle$ 中也可能只有某些 y_i 属于类 B. 我们用表达式 $N^n(A)$ 表示 $\langle x_i \rangle$ 的前 n 项中满足 $x_i \in A$ 的元素的数目，用 $N^n(A \wedge B)$ 表示 $\langle x_i, y_i \rangle$ 的前 n 项中同时满足 $x_i \in A$ 且 $y_i \in B$ 的元素序对的数目，则相对频率 $F^n(A, B)$ 就定义为

$$F^n(A, B) \triangleq N^n(A \wedge B)/N^n(A).$$

因此 Reichenbach 的概率蕴涵的频率解释可以表述如下：若对给定序列 $\langle x_i \rangle, \langle y_i \rangle$，当 $n \to \infty$ 时，$F^n(A, B)$ 趋于极限 p，则 p 称为相对 $\langle x_i \rangle, \langle y_i \rangle$ 的从 A 得到 B 的概率,记作

$$P(A, B) = \lim_{n \to \infty} F^n(A, B) = p.$$

现在我们来证明频率解释模型满足 Reichenbach 的公

理系统。

考虑公理 2.1。假设 $(x_i)\urcorner(x_i\in A)$ 成立，则

$$N^n(A\wedge B)=0=N^n(A),$$

所以 $F^n(A,B)$ 假设未定型 $\frac{0}{0}$ 型，因此我们有

$$P(A,\ B)=\frac{0}{0},$$

即该概率没有确定的值。这个结果表示，若 $p\neq q$，则从 $(x_i)\urcorner(x_i\in A)$ 可推出

$$\left(A\underset{p}{\ni}B\right)\wedge\left(A\underset{q}{\ni}B\right).$$

反之，若 $(\exists x_i)(x_i\in A)$，则存在确定的极限，因为只能有一个极限存在，所以公理 2.1 的另一半也满足。这里须注意，即使只有有穷个元 x_i 属于 A，也仍然存在极限。这时我们把关于最后一个元素的相对频率看作是极限。这种平凡的情况也包括在频率解释中，并且在证明公理 2.1 和下列其他公理时不产生任何困难。

考虑公理 2.2(1)。假设 $A\rightarrow B$ 普遍有效，则根据逻辑蕴涵的含义，对任意 x_i, y_i，若 $x_i\in A$，则 $y_i\in B$。在这种情况下，所有的 $F^n=1$，立即可得 $\left(A\underset{1}{\ni}B\right)$。注意该公理的主联结词不能是等价号，因为存在某些场合，其中有些 $x_i\in A$ 对应 $y_i\notin B$，这时也能得到 $\lim\limits_{n\to\infty}F^n=1$。这样，说明了为何极大概率蕴涵“$\underset{1}{\ni}$”要比逻辑蕴涵“$\rightarrow$”更广泛，这也证明了程度 p 的概率蕴涵表示经典逻辑全称蕴涵的一个概括。

公理 2.2(2) 直接从下列事实得到满足：相对频率是一非负数且总是不大于 1。

为了证明公理 2.3 的满足性，我们首先构造

$$F^n(A,B\vee C)=N^n(A\wedge(B\vee C))/N^n(A),$$

若 $A\to\neg(B\wedge C)$ 普遍有效，这等于（根据频率解释的定义）

$$(N^n(A\wedge B)/N^n(A))+(N^n(A\wedge C)/N^n(A)).$$

因为 B 和 C 相对 A 不相容，即当 $x_i\in A$ 时，对应的 y_i 和 z_i 不能同时分属于 B 和 C，因此我们有

$$F^n(A,B\vee C)=F^n(A,B)+F^n(A,C).$$

该等式在转移到极限过程中始终保持不变，因此对互不相容的事件，我们有

$$P(A,B\vee C)=P(A,B)+P(A,C).$$

考虑公理 2.4，据相对频率的定义，有

$$\begin{aligned}&F^n(A,B\wedge C)\\&=N^n(A\wedge B\wedge C)/N^n(A)\\&=(N^n(A\wedge B)\cdot N^n(A\wedge B\wedge C))/\\&\quad(N^n(A)\cdot N^n(A\wedge B)),\end{aligned}$$

这里假设 $N^n(A\wedge B)\neq 0$，否则易证

$$上式=F^n(A,B)\cdot F^n(A\wedge B,C).$$

该等式在转移到极限过程中始终保持不变．若极限存在，根据极限定义，就有

$$P(A,B\wedge C)=P(A,B)\cdot P(A\wedge B,C).$$

最后证明存在规则也可以从频率解释中推出，因为每一公理表示频率之间的普遍有效关系。这种关系甚至在转移到极限之前也严格保持，所以从这些公理推出的每一概率公式也对应频率之间的普遍有效关系，且这种关系在转移到极限之前也严格有效。每一这样的关系可以写作以下形式

$$f_m^n=r(f_1^n,\cdots,f_{m-1}^n),\tag{2.5}$$

其中 f_i^n 代表下列形式的频率表达式：

$$f_i^n=F^n(A_i,B_i).\tag{2.6}$$

(2.5)和(2.6)的下标表明下列事实：我们在此研究的频率量是属于不同的事件 $A, B, \cdots$. 根据存在规则，r 是一单值函数。当 $n \to \infty$，从有关极限的定律可得：每当 $f_1^n, \cdots, f_{m-1}^n$ 趋于极限 $p_1, \cdots, p_{m-1}$，则 f_m^n 也必然趋于极限 p_m. 换句话说，据存在规则，每当概率 $p_1, \cdots, p_{m-1}$ 存在时，概率 p_m 必然存在。

这样就证明了 Reichenbach 的系统被频率解释模型满足。

显然上述频率解释是针对事件序列的，所以原则上，对单个事件，频率解释是有问题的。Reichenbach 为了把他的频率解释贯彻到底，就提出他的认定理论。他说，假定在一序列中，事件 B 的频率是 5/6. 若有人问，单个事件 B 是否要发生，我们会选择肯定的回答，因为我们如此重复下去，总有 5/6 的可能是对的。那么究竟在什么意义上我们能下此断定呢？回答是，不能把该陈述句看作是一个断定，而是看作一个认定(posit). 这里使用"认定"一词是和博奕论中使用"打赌"一样。当我们赌一匹马的胜负时，我们不是说，根据这样的打赌该马将取胜这句话为真，而是在下赌注时，把这句话当作真的。因此，Reichenbach 对"认定"下了以下定义："认定是我们看作真的陈述句，尽管它的真值是未知的。"([2], p. 373). 当这似乎是我们所能做到的最好选择时，我们才决定作出认定。这里的"最好"具有能用数值解释的意义，它是指当我们重复运用该认定时，它将是最成功的。

认定方法适于用概率陈述句来判定单个事件，它在所有的实际运用中都起重要的作用。例如农民对于未来的天气，医生对于他面临的病人，都要通过对他们认为最有可能出现的事件的假定来做出认定。

Reichenbach 认为，严格地说，只有关于类的概率概念才

是合法的，单个事件的概率可以看作是权宜之计，它必须用类概念来构造. 这个构造就是把单个事件看作是变得越来越窄的类的极限. 这种认定方法已在概率论中根据下列事实证明是正当的： 若我们用子序列的概率 $P(A\wedge S, B)$ 而不是用 $P(A,B)$ 来作为我们认定的根据，我们就能得到更大数目的成功的话，那么如此重复地把主序列划分成子序列，将会导致越来越好的结果，只要这一概率在每一步都在增加. 当该单个事件一直处于越来越窄的类中，这个概率就趋于极限.

§3 作为多值逻辑的概率逻辑

前面我们把概率看作是事件(序列)的性质，但我们也可以把概率看作是关于事件的句子的性质. 通过这样的转换，概率成了命题的信用 (rating) 程度，概率陈述句不再属于对象语言，而是属于元语言. Reichenbach 称这种思想为概率的逻辑解释. 经典逻辑是二值逻辑，把概率看作是一种真值就是要构造一种多值逻辑. Reichenbach 构造的概率逻辑就是真值在[0,1]区间上取连续度量的多值逻辑.

Reichenbach 的概率逻辑主要是相对命题序列而言的. 为了构造恰当的命题序列，我们从下列概率蕴涵式开始讨论.

$$(x_i)(y_i)\left(x_i\in A \underset{p}{\ni} y_i\in B\right), \tag{2.7}$$

(2.7) 实际上表示两个事件序列 $\langle x_i\rangle$, $\langle y_i\rangle$ 分别相对类 A 和类 B 的概率蕴涵关系. 我们现在要考虑关于事件的句子的概率蕴涵关系，所以 (2.7) 可改写为

$$(i)\left(a_i \underset{p}{\ni} b_i\right). \tag{2.8}$$

它实际是对下列概率蕴涵句的缩写

$$a_1 \underset{p}{\supset} b_1,\ a_2 \underset{p}{\supset} b_2, \cdots,$$

即我们考虑的是两个命题序列 $\langle a_i \rangle, \langle b_i \rangle$ 之间的概率蕴涵关系．用 P 记号，我们有

$$P(\langle a_i \rangle, \langle b_i \rangle) = p. \tag{2.9}$$

频率解释根据下列方式构造：在由真命题 a_i 选出的子序列中统计对应的真命题 b_i 的数目(若无特别说明，从现在起概率表达式属于元语言．这是可能的，因为关于元语言的表达式的所有规则严格对应关于对象语言的表达式的规则)．

若把概率看作是一种可以与真值相比较的性质，就必须进一步克服一个障碍．我们知道真值是句子的性质，而概率严格说是两个句子之间的关系，因此我们如下来消除这种困难：用全类 $\langle a_i \vee \neg a_i \rangle$ 来置换 $\langle a_i \rangle$，所以(2.9)可写作

$$P(\langle a_i \vee \neg a_i \rangle, \langle b_i \rangle) = P(\langle b_i \rangle) = p, \tag{2.10}$$

这样出现的概率作为命题序列的性质，与真值相同．

现在我们对 (2.10) 构造一频率解释，这个解释与前一节的频率解释在本质上一致，但这里我们使用的是概率的逻辑解释，所以我们必须对关于事件的句子而不是对事件本身计数，因此可用以下形式写出 (2.10) 的频率解释

$$P(\langle b_i \rangle) = \lim_{n \to \infty} \frac{1}{n} \mathop{\mathbf{N}}_{i=1}^{n} (\sigma(b_i) = 1), \tag{2.11}$$

其中 σ 是古典二值真值指派，$\mathop{\mathbf{N}}\limits_{i=1}^{n} (\sigma(b_i) = 1)$ 表示序列 $\langle b_i \rangle$ 的前 n 项中真命题 b_i 的数目．

对有穷命题序列 $\langle b_i \rangle$，概率的频率解释仍保持，只要我们把 $F^n(\langle b_i \rangle) = \frac{1}{n} \mathop{\mathbf{N}}\limits_{i=1}^{n} (\sigma(b_i) = 1)$ 的值的极限理解为该序列最后一项的相对频率．在这种情况下，一共只有 $n+1$ 个值：$0/n, 1/n, \cdots, n/n$ 对频率是可能的，因此只需 $n+1$ 值

逻辑就够了，无需再涉及连续度量的概率逻辑．此外，有穷序列和无穷序列还有一个重要的区别：对于前者，频率的极限为 1 当且仅当所有的 b_i 皆真，因此相对有穷序列，无须区别全称蕴涵和极大概率蕴涵 $\underset{1}{\ni}$ 之间的差异，但这种情况对无穷序列不成立．

当 n 越来越小时，$n+1$ 值逻辑就向二值逻辑转化．当 $n=1$ 时，就归约为二值逻辑．命题序列也归约为单个命题．因为这时这样的序列的频率只是 1 或 0，所以我们有

$$P(b_1)=1 \text{ 或 } P(b_1)=0.$$

在这种场合，两个可能的概率值和古典真值：真、假相一致，所以我们可以把真假看作当该序列归约为单个命题时产生的概率的极限情况．

下面 Reichenbach 给出具有连续度量的概率逻辑的真值表．为此首先给出以下定义

$$\langle a_i\rangle\vee\langle b_i\rangle\triangleq\langle a_i\vee b_i\rangle,$$
$$\langle a_i\rangle\wedge\langle b_i\rangle\triangleq\langle a_i\wedge b_i\rangle,$$
$$\langle a_i\rangle\rightarrow\langle b_i\rangle\triangleq\langle a_i\rightarrow b_i\rangle,$$
$$\langle a_i\rangle\longleftrightarrow\langle b_i\rangle\triangleq\langle a_i\longleftrightarrow b_i\rangle,$$
$$\neg\langle a_i\rangle\triangleq\langle\neg a_i\rangle,$$
$$P(\langle a_i\rangle,\langle b_i\rangle)\triangleq$$
$$\lim_{n\to\infty}\left(\overset{n}{\underset{i=1}{\mathbf{N}}}(\sigma(a_i\wedge b_i)=1)\Big/\overset{n}{\underset{i=1}{\mathbf{N}}}(\sigma(a_i)=1)\right).$$

表 1　概率逻辑的真值表

一元联结词

$P(\langle a_i\rangle)$	$P(\neg\langle a_i\rangle)$
p	$1-p$

二元联结词和运算

$P(\langle a_i\rangle)$	$P(\langle b_i\rangle)$	$P(\langle a_i\rangle,\langle b_i\rangle)$	$P(\langle a_i\rangle\vee\langle b_i\rangle)$	$P(\langle a_i\rangle\wedge\langle b_i\rangle)$
p	q	u	$p+q-p\cdot u$	$p\cdot u$

$P(\langle a_i\rangle\rightarrow\langle b_i\rangle)$	$P(\langle a_i\rangle\longleftrightarrow\langle b_i\rangle)$	$P(\langle b_i\rangle,\langle a_i\rangle)$
$1-p+p\cdot u$	$1-p-q+2p\cdot u$	$\frac{p\cdot u}{q}$

表 1 包括两个限制条件，

$$1.\ \frac{p+q-1}{p}\leqslant u\leqslant\frac{q}{p},$$

$$2.\ P(\langle a_i\rangle,\langle a_i\rangle)=1,$$

其中条件 2 具有类似前面公理 2.2(1) 的功能。 表中垂直双线左边是变目的真值。显见，作为多值逻辑的概率逻辑和二值逻辑之间存在本质的差异：对后者，除了ㄱ，其它联结词的真值表用两个组元的真值作为变目就能确定复合命题的真值。例如，当 a 的真值为 1，b 的真值为 0 时，$a\wedge b$ 的真值为 0。但在概率逻辑的二元联结词和运算中则不然，两个变目不足以确定复合命题序列的真值，还需要第三个变目. Reichenbach 是用 $P(\langle a_i\rangle)$,$P(\langle b_i\rangle)$ 和 $P(\langle a_i\rangle,\langle b_i\rangle)$ 作为变目的。别的组合可以用这三个值来确定。

易证，上表是古典二值逻辑的概括，即古典二值逻辑可以看作是概率值只取 0 和 1 的特例。在这种情况下，表 1 的头两栏的取值只有 4 种组合，其他复合命题的真值由表 2 给出。

表 2

1	1	1	1	1	1	1	1
1	0	0	1	0	0	0	?
0	1	?	1	0	1	0	0
0	0	?	0	0	1	1	?

若 $p=1$，则据限制条件，$u=q$，所以表 1 的第三栏失去了它的独立地位．这意味表 1 的第三栏已成为前两栏的函数，所以隔开变目和函数的垂直双线就得向左移动一栏．对 $p=0$ 的情况，u 值的不确定性用问号"? "来表示．这样，若令 $p=q=u=1$，则得到析取值 $1+1-1\cdot1=1$，合取值 $1\cdot1=1$，蕴涵值 $1-1+1\cdot1=1$，等等．易证在二值逻辑的所有命题联结词的真值中都能去掉由"?"表示的不确定值．例如，当 $p=0, q=1, u=?$ 时，析取值就是 $0+1-0\cdot?=1$，因为"? "只表示 $[0, 1]$ 中任意的有穷数，它与 0 相乘总是为 0．

Reichenbach 把处理单个命题的概率逻辑称为权逻辑 (logic of weight)，它是通过把命题序列的真值（真值概率）性质虚拟地转移到单个命题而构造起来的．在此我们允许单个命题 a 进入概率函子 (functor) $P(a)=p$，其中数值 p 测度 a 的权．不言而喻，a 的权是根据适当的参考类（如全类）来确定．若用 a,b 置换表 1 中的 $\langle a_i\rangle$ 和 $\langle b_i\rangle$，我们不难得到权逻辑的真值表，从而确定复合命题的权．

§4 归纳法与归纳逻辑

Reichenbach 首先把归纳法分为初级归纳法和高级归纳法．传统的简单枚举法大体相当初级归纳法．关于高级归纳法，他谈到三种常见类型．第一类是 F. Bacon 的三表法和 J. S. Mill 的求因果五种方法．

第二种称为交叉归纳法 (cross induction)，它是对简单枚举法的修正，这个推理可用下列模式来说明．

$$x_1\in B,\ x_2\in B,\ x_3\notin B,\ x_4\in B,\ x_5\notin B,\cdots,$$
$$y_1\in C,\ y_2\notin C,\ y_3\in C,\ y_4\in\ \ ,\ y_5\notin C,\cdots,$$
$$\vdots$$

⋮

$$z_1 \in S,\ z_2 \in S,\ z_3 \in S,\ z_4 \in S,\ z_5 \in S, \cdots.$$

对最上的横序列，归纳推理能得到下列结果：正负事例都出现，对最下的横序列，归纳推理的结论是：总是出现正事例。然后再在垂直方向做归纳推理，这时把每一横序列作为一个元素，因为上面所有别的序列都出现正负事例，所以我们推出，最下的横序列只要是充分连续的，它也会出现负事例。这样，我们通过交叉归纳法，消去了一个归纳推理。这种推理模式经常使用。例如，虽然碳还没有熔化，但在更高的温度下它将熔化，因为其他所有的物质都如此。

第三种是解释归纳法 (explanatory induction)：从一定的观察材料到假说或理论的推理。根据一定的观察材料，往往能提出多个解释这些材料的假说，运用归纳法和 Bayes 定理，可以分别求得这些假说的概率，从而进一步选择最合适的假说和理论。

以上提出的各种形式的归纳法包括一个共同的特征：推理不仅建立在新的观察材料之上，而且建立在前面的归纳推理的成果之上，因此归纳法的正当性建立在一种把许多归纳推理结合成网络推理的方法之上。

所有不具有枚举归纳法形式的归纳推理必须用概率演算的定理来解释。而概率演算的公理化系统导致下列结果：若假定频率解释，所有的概率推理都可以归约为演绎推理和枚举归纳法。在概率蕴涵句之间以及与别的公式之间的关系和推演需要用演绎推理。确定概率蕴涵句需要用枚举归纳法，因为一般说，从相对频率，只有经过认定一修正一认定的方法才能得到极限，而这种方法本质上就是用概率论语言刻划的枚举归纳法。所以可以得出结论，所有的归纳推理都可以归约为枚举归纳法，当然这并不等于说所有的归纳推理都可

以直接转换为枚举法。这种归约只能间接地把概率公理归约为频率解释，因为在频率解释下，概率就是极限频率，而根据极限频率的定义，所谓 p 是相对频率 $F^1, F^2, \cdots$ 的极限，就是说，对任意 $\varepsilon > 0$，总存在 N，使得对所有 $n > N$，都有

$$|p - \varepsilon| \leqslant F^n \leqslant p + \varepsilon.$$

我们知道，在一般情况下序列的元素的出现无一定规则。对这样的序列求极限，只有通过考察相对频率来认定极限频率，即通过考察 $F^1, F^2, \cdots, F^N$ 认定极限频率随 $n \to \infty$ 将趋于其中某个 $F^i (i \leqslant N)$，这里的 F^i 可以是这些相对频率的平均数，也可以是在这些频率中出现次数最多的。显然，这样做出的认定未必一定正确，于是需要再考察 $F^{N+1}, F^{N+2}, \cdots, F^{N+m}$，重新认定某个 F^i 是极限频率。若类 B 相对参考类 A 的概率 p 存在，则极限频率就存在，根据极限频率的定义，对任给 $\varepsilon > 0$，总存在自然数 N^* 使得 $F^{N^*+1}(A, B)$，$F^{N^*+2}(A, B), \cdots$，都落在区间

$$[\,|F^{N^*}(A,B) - \varepsilon|,\ F^{N^*}(A,B) + \varepsilon]$$

中，从而 $F^{N^*}(A,B)$ 就是我们要求的极限频率。因此 Reichenbach 认为，若极限频率存在，则通过认定一修正一认定的过程，就一定能从相对频率中把它求出来。这样的过程就是归纳推理，这样的规则就是归纳规则。归纳逻辑就是包含归纳规则和所有演绎逻辑的推演规则和公理。因此在 Reichenbach 看来，归纳逻辑和概率逻辑是有区别的，后者是演绎逻辑，是古典二值逻辑的概括，而前者是演绎逻辑的所有规则和公理，包括概率逻辑的所有规则，外加上述归纳规则。

§5 Reichenbach 对我们的启发

如果说 Keynes 的概率演算是建立在古典命题演算的基

础上，那么 Reichenbach 的工作是建立在谓词演算的基础上的，所以我们遵循 Reichenbach 的思想，试图得到以下概括形式。

给定一阶语言 $\mathscr{L}$，我们添加概率蕴涵号 $\underset{p}{\ni}$，其中 $p\in[0,1]$。注意：实际上我们添加了连续统多个概率蕴涵号。

形式规则增加：

若 $\varphi(x)$，$\psi(x)$ 是公式，则 $(x)\left(\varphi(x)\underset{p}{\ni}\psi(x)\right)$ 也是（注意：$\varphi(x)$，$\psi(x)$ 中还可以有别的自由变元）。

根据 Reichenbach 在本章§1的思想，我们有如下公理系统：除了谓词演算的公理和规则，我们还有

公理 2.6（单一性）

$$(x)(\varphi(x)\underset{p}{\ni}\psi(x))\wedge(x)(\varphi(x)\underset{q}{\ni}\psi(x))\longleftrightarrow(x)\neg\varphi(x),$$

其中 $p\neq q$。

公理 2.7（正规性）

(1) $(x)(\varphi(x)\to\psi(x))\to(x)\left(\varphi(x)\underset{1}{\ni}\psi(x)\right)$.

(2) $(x)\left(\varphi(x)\underset{p}{\ni}\psi(x)\right)$，其中 $p\geqslant 0$.

公理 2.8（加法公理）

(1) $(x)(\varphi(x)\to\neg(\psi(x)\wedge\theta(x)))\to$

$\left((x)\left(\varphi(x)\underset{p}{\ni}\psi(x)\right)\wedge(x)\left(\varphi(x)\underset{q}{\ni}\theta(x)\right)\to\right.$

$\left.(x)\left(\varphi(x)\underset{p+q}{\ni}\psi(x)\vee\theta(x)\right)\right)$.

(2) $(x)(\varphi(x)\to\neg(\psi(x)\wedge\theta(x)))\to$

$\left((x)\left(\varphi(x)\underset{p}{\ni}\psi(x)\right)\wedge(x)\left(\varphi(x)\underset{q}{\ni}\psi(x)\vee\theta(x)\right)\to\right.$

$$(x)\left(\varphi(x) \underset{q-p}{\supset} \theta(x)\right)\right).$$

公理 2.9 （乘法公理）

(1) $(x)\left(\varphi(x) \underset{p}{\supset} \psi(x)\right) \wedge (x)\left(\varphi(x) \wedge \psi(x) \underset{q}{\supset} \theta(x)\right)$

$\rightarrow (x)\left(\varphi(x) \underset{p \cdot q}{\supset} \psi(x) \wedge \theta(x)\right).$

(2) $(x)\left(\varphi(x) \underset{p}{\supset} \psi(x) \wedge \theta(x)\right) \wedge (x)\left(\varphi(x) \underset{q}{\supset} \psi(x)\right)$

$\rightarrow (x)\left(\varphi(x) \wedge \psi(x) \underset{p/q}{\supset} \theta(x)\right)$，其中 $q \neq 0$。

(3) $(x)\left(\varphi(x) \underset{p}{\supset} \psi(x) \wedge \theta(x)\right) \wedge (x)\left(\varphi(x) \wedge\right.$

$\left.\psi(x) \underset{q}{\supset} \theta(x)\right) \rightarrow (x)\left(\varphi(x) \underset{p/q}{\supset} \psi(x)\right)$，其中 $q \neq 0$.

注意，用此公理系统和第一章§4我们建立的公理系统相比较，它们在形式上并无太大的区别。此外，和第一章§4一样，我们利用一组公理来描述加法公理和乘法公理，这是为了增加系统的推演力，保证能推出诸如 Bayes 定理那样的重要定理。

下面我们推演几个定理以说明上述系统的特征。

定理 2.1 $\left(\varphi \underset{1}{\supset} \psi \vee \neg\psi\right).$

证明：

据公理 2.7(1) 等，我们得

$$(\varphi \rightarrow \psi \vee \neg\psi) \rightarrow \left(\varphi \underset{1}{\supset} \psi \vee \neg\psi\right).$$

根据分离规则，我们有

$$\left(\varphi \underset{1}{\supset} \psi \vee \neg\psi\right).$$

定理得证.

定理 2.2

$$(x)\left(\varphi(x) \underset{p}{\supset} \psi(x)\right) \longleftrightarrow (x)\left(\varphi(x) \underset{1-p}{\supset} \neg\psi(x)\right).$$

证明：根据公理 2.8(2) 和分离规则，可得

$$(x)(\varphi(x) \to \neg(\psi(x) \wedge \neg\psi(x))) \to$$

$$\left((x)\left(\varphi(x) \underset{p}{\supset} \psi(x)\right) \wedge (x)\left(\varphi(x) \underset{q}{\supset} \psi(x) \vee \neg\psi(x)\right)\right.$$

$$\left.\to (x)\left(\varphi(x) \underset{q-p}{\supset} \neg\psi(x)\right)\right),$$

$$(x)\left(\varphi(x) \underset{q}{\supset} \psi(x) \vee \neg\psi(x)\right) \to \left((x)\left(\varphi(x) \underset{p}{\supset} \psi(x)\right)\right.$$

$$\left.\to (x)\left(\varphi(x) \underset{q-p}{\supset} \neg\psi(x)\right)\right).$$

令 $q = 1$，且由定理 2.1，可得

$$(x)\left(\varphi(x) \underset{p}{\supset} \psi(x)\right) \to (x)\left(\varphi(x) \underset{1-p}{\supset} \neg\psi(x)\right),$$

$$(x)\left(\varphi(x) \underset{1-p}{\supset} \neg\psi(x)\right) \to (x)\left(\varphi(x) \underset{p}{\supset} \psi(x)\right),$$

定理得证. 我们可将定理 2.2 与第一章定理 1.4，定理 1.20 相比较.

根据 Reichenbach 在本章 § 2 的思想，我们来构造概括 Reichenbach 频率解释的语义学.

令一阶语言 $\mathscr{L} = \langle R_i, F_j, a_k \rangle_{i\in I, j\in J, k\in K}$，我们的模型是下列形式的四元组的结构：

$$\mathfrak{U} = \langle A, R_i^{\mathfrak{U}}, F_j^{\mathfrak{U}}, a_k^{\mathfrak{U}} \rangle_{i\in I, j\in J, k\in K},$$

其中 A 是可数非空论域 $A=\langle a_i : i=1,2,\cdots\rangle$，$R_i^n$，$F_j^n$，$a_k^n$ 如通常定义。

此外，除了满足关系通常的递归定义外，还增加

$$\mathfrak{A}\models(x)\left(\varphi(x)\underset{p}{\supset}\psi(x)\right)\Longleftrightarrow\lim_{n\to\infty}(|\{a\in A_n:\mathfrak{A}\models\varphi(x)\wedge\psi(x)[a]\}|/|\{a\in A_n:\mathfrak{A}\models\varphi(x)[a]\}|)=p,$$

其中 $A_n=\{a_i : i=1,2,\cdots,n\}$，$|A|$ 表示 A 的基数(即 A 的元素的个数)，且 $|\{a\in A_n:\mathfrak{A}\models\varphi(x)[a]\}|\neq 0$。

显然这样的模型满足上述公理系统。

从上述结果(包括我们对 Reichenbach 工作的概括)可见，Reichenbach 的工作是有缺陷的，主要集中在他的频率解释理论。虽然他采用了认定理论进行补救，但终究被人们认为是特设性的。此外把概率只看作是相对频率的极限，这个限制非常大，例如他考虑的序列不仅要求可数而且还需要一一对应，极限必须存在等等。直到本世纪 70 年代才真正解决了这些问题，这就是本书最后一章所要介绍的内容。

第三章 Carnap 的归纳逻辑

R. Carnap 的归纳逻辑继承了 Keynes 的关于归纳概率是证据和假说之间的一种逻辑关系的思想，并且批判了 Reichenbach 的频率解释理论. 他把概率概念分成两类：概率 1 和概率 2. 概率 2 是(以 Reichenbach 为代表)频率概念. 当它用于统计研究时，是一种重要的科学概念，但作为归纳逻辑的基本概念是不恰当的. 概率 1 就是他的确证函数，它定义了观察命题对假说命题的确证度. 概率 1 是逻辑概念，因为给定观察证据 e，确定假说 h 是否被确证，以及在什么程度上确证是一逻辑问题. 虽然，这里的句子 e 和 h 自身涉及到经验事实，但一旦 h 和 e 给定，上述问题就只要求我们能够分析它们的意义，以及建立在这些意义之上的逻辑关系.

Carnap 的归纳逻辑(他认为归纳逻辑就是对归纳推理提供推演规则的逻辑概率理论)理论十分庞大. 50 年代初的工作主要包括构造定量归纳逻辑、可比较归纳逻辑、归纳相干理论、归纳估计理论和归纳方法的连续统理论. 60 年代后，他对概率的解释向主观主义退却. 在技术上，运用模型论、集合论概率公理概括了他早期的归纳逻辑，使之更具一般性. 我们在下面两节分别讨论他的定量归纳逻辑和他的归纳方法的连续统理论，其余请参见 [3]，[4] 和 [5].

§1 构造定量归纳逻辑

Carnap 的语言系统(经过语义解释了的语言，实际就是

指模型)$\mathscr{L}$ 包括无穷系统 $\mathscr{L}^{\pi}_{\infty}$ 和有穷系统序列 $\mathscr{L}^{\pi}_{N}$，后者的 N 跑遍 (range over) 所有的正整数，因此形成了语言系统的一无穷序列：$\mathscr{L}^{\pi}_{1}$，$\mathscr{L}^{\pi}_{2}$，…，其中 $\mathscr{L}^{\pi}_{N+1}$ 除了比 $\mathscr{L}^{\pi}_{N}$ 多一个体常元 a_{N+1} 之外，与 $\mathscr{L}^{\pi}_{N}$ 的符号完全相同。

$\mathscr{L}$ 包括下列经语义解释了的非逻辑符号：

(1) 个体常元：$a_1, a_2, \cdots, a_N$ (对 $\mathscr{L}^{\pi}_{N}$)，$a_1, a_2, \cdots$ (对 $\mathscr{L}^{\pi}_{\infty}$)。

(2) π 个(有穷个)一元初始谓词：$P_1, P_2, \cdots, P_\pi$。

上述个体常元和一元初始谓词必须满足下列独立性和完全性要求。

(i) 独立性要求：个体常元必须是指称不同的、分立的个体，即若 a, b 是 $\mathscr{L}$ 的任意两个体，P 是任一谓词，则形如 $P(b) \wedge \neg P(a)$ 不是自相矛盾的。初始谓词指称在逻辑上是互相独立的属性，即若 P_i，P_j 是 $\mathscr{L}$ 的两谓词，a 是任一个体，则形如 $P_i(a) \wedge \neg P_j(a)$ 不是自相矛盾的。

(ii) 完全性要求：任意两个体只在有穷多方面不同。初始谓词足以表达我们给定的论域中的个体表现出来的所有定性的性质。

$\mathscr{L}$ 的逻辑符号是带等词的一阶谓词演算的符号，其中 $\neg, \vee, \wedge$ 是初始联结词，其他的定义和形成规则如通常。此外还有"m"(测度函数)和"c"(确证度函数)，它们不是为了指称对象语言的表达式，而是用于元语言中。

下面我们给出一系列的定义和定理，用于构造定量归纳逻辑的语义规则。这些规则确定 $\mathscr{L}$ 中所有句子的意义。

定义 3.1 令 $i \in S(\mathscr{L})$ (其中 $S(\mathscr{L})$ 表示 $\mathscr{L}$ 的句子集，下同)，称 i 在 $\mathscr{L}$ 中真，若 i 满足下列条件之一。

(1) i 是原子句 $P(a)$(P 是 $\mathscr{L}$ 的任一一元初始谓词，a 是任一个体)，且 a 具有 P 指称的属性。

(2) i 是"$a=a$".

(3) i 是"t"("t"是古典命题逻辑的重言式或它的代入特例).

(4) i 是 $\neg j$, 且 j 在 $\mathscr{L}$ 中不真.

(5) i 是 $j \vee k$, 且 j 在 $\mathscr{L}$ 中真或 k 在 $\mathscr{L}$ 中真.

(6) i 是 $j \wedge k$, 且 j 在 $\mathscr{L}$ 中真并且 k 在 $\mathscr{L}$ 中真.

(7) i 是 $(x)A$, 且对 A 的每一代入特例 A^* (A^* 是用 $\mathscr{L}$ 的个体常元替换 A 中自由出现的 x 得到的公式), A^* 在 $\mathscr{L}$ 中为真.

定义 3.2 (从定义 3.2 到定理 3.16 我们用 $\mathscr{L}_N$ 表示 $\mathscr{L}_N^\pi$,$\mathscr{L}_\infty$ 表示 $\mathscr{L}_\infty^\pi$)

(1) 称 SD 是 $\mathscr{L}_N$ 的一状态描述 (state description), 若 SD 是一合取式, 使得它从 $\mathscr{L}_N$ 的每一基本句子对(由 $\mathscr{L}_N$ 的任意原子句和它的否定组成)中仅取一个作为合取肢, 且不取别的任何句子,其中这些合取肢的排列按词典顺序.

(2) 称 SD 是 $\mathscr{L}_\infty$ 的一状态描述,若 SD 是一集合,使得它从 $\mathscr{L}_\infty$ 的每一基本句子对中仅取一个为元素,且不含别的任何句子.

定义 3.3 设任意 $i \in S(\mathscr{L})$, SD 是 $\mathscr{L}$ 的任一状态描述, i 在 $\mathscr{L}$ 的 SD 中成立(简写为 $\mathscr{L}, SD \models i$), 若满足下列条件之一.

(1) i 是一原子句,且 i 是 SD 的某个合取肢(对 $\mathscr{L}_N$), 或 $i \in SD$ (对 $\mathscr{L}_\infty$).

(2) i 具有这样的形式 $a=a$.

(3) i 是"t".

(4) i 是 $\neg j$, 且 $\mathscr{L}, SD \not\models j$ (j 在 $\mathscr{L}$ 的 SD 中不成立).

(5) i 是 $j \vee k$, 且 $\mathscr{L}, SD \models j$ 或 $\mathscr{L}, SD \models k$.

(6) i 是 $j\wedge k$，且 $\mathscr{L}$,SD$\models j$ 并且 $\mathscr{L}$,SD$\models k$.

(7) i 是 $(x)A$,且 $\mathscr{L}$,SD$\models A^*$(A^* 定义如定义 3.1).

定义 3.4 若 $S\subseteq$S($\mathscr{L}$)，称 $\mathscr{L}$，SD$\models S$，若对每一 $i\in S$，都有 $\mathscr{L}$,SD$\models i$.

定义 3.5 令 $i\in$S($\mathscr{L}$),$S\subseteq$S($\mathscr{L}$),

(1) $\mathscr{L}$中 i 的适域 (range) RA(i) = {SD ∈ VSD($\mathscr{L}$): $\mathscr{L}$,SD$\models i$}，其中 VSD($\mathscr{L}$) = {SD: SD 是 $\mathscr{L}$ 的状态描述}(在不引起误解的地方,可简写为 VSD).

(2) $\mathscr{L}$ 中 S 的适域 RA(S) = {SD ∈ VSD($\mathscr{L}$): $\mathscr{L}$, SD$\models S$}.

定理 3.1 (适域定理) 令 i,S 定义如前.

(1) 若 i 是一原子句,则 RA(i) = {SD: i 是 SD 的合取肢(对 $\mathscr{L}_N$)，或 $i\in$SD (对 $\mathscr{L}_\infty$)}.

(2) 若 i 是 $a=a$，则 RA(i) = VSD.

(3) 若 i 是 $a=b$(其中 a,b 是 $\mathscr{L}$ 的两不同个体常元)，则 RA(i) = $\varnothing$.

(4) 若 i 是"t"，则 RA(i) = VSD.

(5) 若 i 是 $\neg j$，则 RA(i) = VSD − RA(j).

(6) 若 i 是 $j\vee k$，则 RA(i) = RA(j)∪RA(k).

(7) 若 i 是 $j\wedge k$，则 RA(i) = RA(j)∩RA(k).

(8) 若 i 是 $(x)A$，则

$$\mathrm{RA}(i)=\bigcap_{j\in J}\mathrm{RA}(j),$$

其中 J = {j: j 是 $\mathscr{L}$ 中 A 的代入特例}.

(9) 若 $S\neq\varnothing$，则

$$\mathrm{RA}(S)=\bigcap_{j\in S}\mathrm{RA}(j),$$

定理 3.2 令 i 是 $\mathscr{L}_N$ 的任一非一般句（非一般句就是无量词句），因而也是 $\mathscr{L}_{N+1}$ 的非一般句．令 SD 和 RA(i) 分别是 $\mathscr{L}_N^*$ 的状态描述和 i 的适域，SD′ 和 RA′(i)是 $\mathscr{L}_{N+1}$ 的状态描述和 i 的适域，则 RA′(i) = {SD′：对任意 SD ∈ RA(i)，SD 是 SD′ 的子合取}．

证明：因为 i 是非一般句，所以它是从定理 3.1(1)，(2)，(3)，(4) 提到的那些简单句通过 ┐，∧，∨ 的布尔组合得到的，根据 SD 和 SD′ 的构造，对 i 的复杂度进行归纳，显证该定理．

定理 3.3 令 i 是 $\mathscr{L}_N$ 的任一非一般句，因而也是 $\mathscr{L}_\infty$ 的非一般句．再令 SD，RA(i) 是 $\mathscr{L}_N$ 的状态描述和 i 的适域，SD′ 和 RA(i)′ 分别是 $\mathscr{L}_\infty$ 的状态描述和 i 的适域，则 RA′(i) = {SD′：对每一 SD ∈ RA(i)，SD 的每一合取肢属于 SD′}．

证明类似定理 3.2．此外可考虑，上述问题为何不能包括 $(x)A$，反例请见定理 3.9 后面的说明．

定义 3.6 令 $i, j_1, j_2, \cdots, j_n \in S(\mathscr{L})$，则有

(1) i（在 $\mathscr{L}$ 中）逻辑真，记作$(\mathscr{L})\models i$，若 RA(i)（在 $\mathscr{L}$ 中）= VSD($\mathscr{L}$)．

(2) i（在 $\mathscr{L}$ 中）逻辑假，记作 $(\mathscr{L})\not\models i$，若 RA($i$)（在$\mathscr{L}$中）= ∅．

(3)（在$\mathscr{L}$中）i 逻辑蕴涵 j，若 $RA(i)\subseteq RA(j)$．

(4)（在$\mathscr{L}$中）i 逻辑等价 j，若 RA(i) = RA(j)．

(5)（在$\mathscr{L}$中）$j_1, j_2, \cdots, j_n$ ($n \geqslant 2$) 逻辑相析（L-disjuct），若 $RA(j_1)\cup RA(j_2)\cup\cdots\cup RA(j_n) = VSD(\mathscr{L})$．

(6)（在$\mathscr{L}$中）i 和 j 逻辑不相容，若 $RA(i)\cap RA(j) = \varnothing$．

定理 3.4 令 $i, j_1, j_2, \cdots, j_n \in S(\mathscr{L})$，则有

(1) i 逻辑假 $\Leftrightarrow \models \neg i$.

(2) i 逻辑蕴涵 $j \Leftrightarrow \models i \rightarrow j$.

(3) i 逻辑等价 $j \Leftrightarrow \models i \leftrightarrow j$.

(4) $i_1, i_2, \cdots, i_n$ 逻辑相析 $\Leftrightarrow \models i_1 \vee i_2 \vee \cdots \vee i_n$.

(5) i 与 j 逻辑不相容 $\Leftrightarrow \models \neg(i \wedge j)$.

定义 3.7 令 $i \in S(\mathscr{L})$，称 i 在 $\mathscr{L}$ 中是事实的（factual），若 i 在 $\mathscr{L}$ 中既不逻辑真又不逻辑假.

定理 3.5 令 $i \in S(\mathscr{L}), S \subseteq S(\mathscr{L})$，$i$（或 S）是事实的 $\Leftrightarrow RA(i)$（或 $RA(S)$）既不等于 $VSD(\mathscr{L})$ 也不等于 $\varnothing$.

定理 3.6 令 i 规定如上，i 是事实的，若 i 满足下列条件之一，

(1) i 是一基本句(原子句或它的否定).

(2) i 是 $\mathscr{L}_N$ 的状态描述.

定理 3.7 令 i 和 j 是 $\mathscr{L}_N$ 的非一般句，因而也是 $\mathscr{L}_{N+1}$ 的非一般句，则有

(1) $\mathscr{L}_{N+m} \models i$ (i 在 $\mathscr{L}_{N+m}$ 中逻辑真) $\Leftrightarrow \mathscr{L}_N \models i$.

(2) $\mathscr{L}_{N+m} \not\models i \Leftrightarrow \mathscr{L}_N \not\models i$.

(3) i 和 j 在 $\mathscr{L}_{N+m}$ 中逻辑蕴涵(逻辑等值，逻辑相析，逻辑不相容) $\Leftrightarrow$ i 和 j 在 $\mathscr{L}_N$ 中逻辑蕴涵(逻辑等值，逻辑相析，逻辑不相容).

(4) i 在 $\mathscr{L}_{N+m}$ 中是事实的 $\Leftrightarrow$ i 在 $\mathscr{L}_N$ 中是事实的.

证明：我们只证明(1). 只须证 $\mathscr{L}_{N+1} \models i \Leftrightarrow \mathscr{L}_N \models i$ 即可，"$\Rightarrow$"显然. 现证"$\Leftarrow$"，令 $SD, RA(i)$ 是 $\mathscr{L}_N$ 的状态描述和 i 的适域，SD' 和 $RA'(i)$ 是 $\mathscr{L}_{N+1}$ 的状态描述和 i 的适域. 若 $\mathscr{L}_N \models i$，则 $RA(i) = VSD(\mathscr{L})$. 据定理 3.2，$RA'(i) = \{SD'$: 对任意 $SD \in VSD(\mathscr{L})$，SD 是 SD' 的子合取$\}$. 显然这时 $RA'(i) = VSD(\mathscr{L}_{N+1})$，否则存在一 $SD^* \in RA'(i)$，使得 i 在 SD^* 上不成立. 因为 i 是 $\mathscr{L}_N$ 的句

子，所以对 SD^*，存在 $\mathscr{L}_N$ 的一状态描述 SD^+，使得它是 SD^* 的子合取且 $\mathscr{L}_N$，$SD^+ \nvDash i$，这与 $RA(i) = VSD(\mathscr{L})$ 矛盾.

定理 3.8 令 i 和 j 是 $\mathscr{L}_N$ 中的非一般句，因而也是 $\mathscr{L}_\infty$ 的非一般句，

(1) $\mathscr{L}_\infty \models i \Longleftrightarrow \mathscr{L}_N \models i$，（证明据定理 3.3）.

(2)—(4) 与定理 3.7 的 (2)—(4) 相同，只是在 $\mathscr{L}_{N+m}$ 之处换成 $\mathscr{L}_\infty$.

定理 3.9 令 i 和 j 是 $\mathscr{L}$ 的任意非一般句，则有

(1) 若 i 在任意系统中逻辑真（逻辑假，事实的），则它在它出现的每一系统中逻辑真（逻辑假，事实的）.

(2) 若在任意系统中，i 和 j 逻辑蕴涵（逻辑等值，逻辑相析，逻辑不相容），则在 i 和 j 出现的任意系统中也同样.

注意，定理 3.9 并不对所有句子成立. 在含有"="的一般句中，很容易找到反例. 例如，令 i 是"$(x)(y)(z)(x = y \vee x = z \vee y = z)$"，$j$ 是"$(\exists x)(\exists y)(x \neq y)$"，即 i 说至多有两个体，j 说至少有两个体. 因为 i 和 j 不包含任何个体常元，所以它在所有的系统中都出现，但是 i 只在 $\mathscr{L}_1$，$\mathscr{L}_2$ 中逻辑真，而在其它系统中逻辑假，j 只在 $\mathscr{L}_1$ 中逻辑假，而在其它系统中逻辑真. 因此，$i \wedge j$ 只在 $\mathscr{L}_2$ 中逻辑真，而在其它系统中逻辑假.

定理 3.10 令 $i, j \in S(\mathscr{L})$，则有

(1) 若 $\mathscr{L}_\infty \models i$，则存在一 m，使得对每一 $\mathscr{L}_N(N \geqslant m)$，$\mathscr{L}_N \models i$.

(2) 相对 i，j，对定理 3.9 的 (2)，(3) 提到的概念，类似 (1) 成立.

证明：我们只证明 (1)：若对任意 m，存在 $\mathscr{L}_N(N \geqslant m)$，$\mathscr{L}_N \nvDash i$，则存在 $\mathscr{L}_N$ 的状态描述 SD 使得 $\mathscr{L}_N$，$SD \nvDash i$. 对 SD 做 $\mathscr{L}_\infty$ 的扩张到 SD′（即若 j 是 SD 的合取肢，则

$i \in SD'$，且从不包含 SD 的合取肢的 $\mathscr{L}_\infty$ 的每一基本句子对中仅取一元放入 SD'），仍有 $\mathscr{L}_\infty$，$SD' \nvDash i$，与题设矛盾。注意，定理 3.10 的逆并不一般成立。

定义 3.8 令 $i \in S(\mathscr{L})$，称 i 是重言的（tautologous），若 i 从组元 $j_1, j_2, \cdots, j_n (n \geqslant 1)$ 中根据联结词构造出来，使得 i 满足下列条件：

（1）每一 j 组元不是别的句子的否定、析取和合取，但它可以是别的任何形式的句子，包括全称形式。

（2）对这些组元的每一次可能的真值赋值（$\boldsymbol{T}$ 和 $\boldsymbol{F}$），根据上述联结词通常的真值表确定的 i 的真值总是取真值 $\boldsymbol{T}$（注意，"i" 总是有真值 $\boldsymbol{T}$）。

定义 3.9 在 $\mathscr{L}_N$ 中，我们称 SD 同构于 SD'，若在 SD 和 SD' 之间存在一个一一对应的常元置换（permutation）函数 $\boldsymbol{C}: T \to T$，其中 $T = \{a_1, a_2, \cdots, a_N\}$，简称 SD 和 SD' 之间存在一 $\boldsymbol{C}$ 置换。

定义 3.10 称 ST(SD) 是 $\mathscr{L}_N$ 中对应 SD 的结构描述，若 $ST(SD) = \bigvee\{x \in VSD(\mathscr{L}_N): x$ 与 SD 同构$\}$。

显然，对 $\mathscr{L}_N$ 的状态描述 SD，与 SD 同构的状态描述有 $N!$ 个。因为从 $a_1, a_2, \cdots, a_N$ 中任选一个元素作为第一个有 N 种可能，然后从其余的 $N-1$ 个元素中任选一个元素作为第二个有 $N-1$ 种可能，…，因此 $N(N-1)(N-2)\cdots 1 = N!$

在做了以上准备工作之后，我们现在可以着手构造 Carnap 的定量归纳逻辑。

定义 3.11 称 m 是相对 $\mathscr{L}_N$ 中状态描述的一正则测度（regular measure）函数（简称测度函数），若 m 满足下列两条件：

（1）对 $\mathscr{L}_N$ 中每一状态描述 SD，$m(SD) > 0$（即

$m(\mathrm{SD})$ 是一正实数)。

(2) $\sum_{\mathrm{SD}\in \mathrm{VSD}(\mathscr{L}_N)} m(\mathrm{SD})=1$。

定义 3.12 设m是 $\mathscr{L}_N$ 的状态描述的测度函数，我们用下列方式把m扩张到 $\mathscr{L}_N$ 的关于句子的测度函数：

(1) 对 $\mathscr{L}_N$ 中任意逻辑假的句子 $j, m(j)=0$。

(2) 对 $\mathscr{L}_N$ 中任意非逻辑假的句子 j，

$$m(j)=\sum_{\mathrm{SD}\in \mathrm{RA}(j)} m(\mathrm{SD}).$$

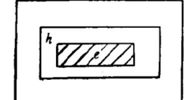
定义 3.13 称 c 是 $\mathscr{L}_N$ 的正则确证函数(简称确证函数)，若m是 $\mathscr{L}_N$ 的关于句子的测度函数，则对任意 $h, e\in S(\mathscr{L}_N)$，

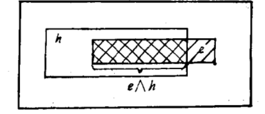
$$c(h,e)=\begin{cases} m(e\wedge h)/m(e) & \text{若 } m(e)\neq 0, \\ \text{没有值} & \text{否则.}\end{cases}$$

根据定义 3.13 定义的确证函数，也称为依赖m的确证函数。

定义 3.13 帮助我们进一步看清了演绎逻辑和归纳逻辑的相似与区别。前者的基本概念是逻辑蕴涵，因此演绎逻辑的基本句具有这样的形式："e 逻辑蕴涵 h"，$\models e\to h \Leftrightarrow \mathrm{RA}(e)\subseteq \mathrm{RA}(h)$。而后者的基本概念是确证度，因此归纳逻辑的基本句具有这样的形式："$c(h,e)=r$"，而从定义 3.13 知 $m(e\wedge h)/m(e)=r$（若 $m(e)\neq 0$）。我们可以把 $m(e)$ 看作是指派给 $\mathrm{RA}(e)$ 的测度，$m(e\wedge h)$ 看作是指派给 $\mathrm{RA}(e\wedge h)(=\mathrm{RA}(e)\cap \mathrm{RA}(h))$ 的测度。换句话说，$\mathrm{RA}(e\wedge h)$ 是 $\mathrm{RA}(e)$ 包含在 $\mathrm{RA}(h)$ 中的部分。例如，"$c(h,e)=3/4$" 述说的是 $\mathrm{RA}(e)$ 的 3/4 包含在 $\mathrm{RA}(h)$ 中。下页图说明上述关系，其中图中面积表示句子的适域。

这样，演绎逻辑和归纳逻辑都研究句子适域之间的关系。而句子的适域是与事实不相干的，只依赖于它自身的意义，这种意义只是由我们讨论中的语言系统的语义规则确定的。若





演绎逻辑
"e逻辑蕴涵 h"表示 e 的适域完全包含在 h 的适域中.

归纳逻辑
"$c(h,e)=3/4$"表示 e 的适域的 3/4 包含在 h 的适域中.

给定这样的语言和规则，就能在这两种逻辑中建立起适域间的关系，无需任何事实(即语言外的、偶然的事实)的知识，这就是为何 Carnap 称他自己的概率概念是逻辑概念的理由.

定理 3.11 令 c 是 $\mathscr{L}_N$ 的确证函数，则存在 $\mathscr{L}_N$ 的关于句子的测度函数 m 使得下式成立.

(1) 对任意 $h,e\in S(\mathscr{L}_N)$，其中 $m(e)\neq 0$，则 $c(h,e)=m(e\wedge h)/m(e)$.

(2) c 对任意 $h,e\in S(\mathscr{L}_N)$ 有值 $\Leftrightarrow m(e)\neq 0\Leftrightarrow \mathscr{L}_N\nvDash \neg e$.

定义 3.14 令 $m^1, m^2,\cdots$ 分别是 $\mathscr{L}_1,\mathscr{L}_2,\cdots$ 的关于句子的测度函数序列，称 m 是对应该序列的 $\mathscr{L}_\infty$ 的关于句子的测度函数，若对任一 $i\in S(\mathscr{L}_\infty)$，$m(i)=\lim\limits_{N\to\infty}m^N(i)$. 若极限不存在，则 m 对 i 没有值.

定义 3.15 令 $c^1,c^2,\cdots$ 分别是 $\mathscr{L}_1$，$\mathscr{L}_2,\cdots$ 的确证函数，称 c 是对应该序列的 $\mathscr{L}_\infty$ 的关于句子的确证函数，若对任意 $h,e\in S(\mathscr{L}_\infty)$，$c(h,e)=\lim\limits_{N\to\infty}c^N(h,e)$. 若极限不存在，则 c 对 h,e 没有值.

定理 3.12 令 m^N，$N=1,2,3,\cdots$ 是测度函数序列，m^∞ 是对应的 $\mathscr{L}_\infty$ 的关于句子的测度函数(当不考虑具体情况时，我们用 m 统一表示 m^N 和 m^∞)，则有

(1) 若 $\models i \longleftrightarrow j$，则 $m(i)=m(j)$。

(2) 若 $\models \neg i$，则 $m(i)=0$。

(3) 若 $\models i$，则 $m(i)=1$。

(4) $0 \leqslant m(i) \leqslant 1$。

(5) 若 $\models i \rightarrow j$，则 $m(i) \leqslant m(j)$。

(6) $m(i \vee j)=m(i)+m(j)-m(i \wedge j)$。

(7) 若 $m(i \wedge j)=0$，从而 $\models \neg(i \wedge j)$，则 $m(i \vee j)=m(i)+m(j)$。

(8) $m(\neg i)=1-m(i)$。

(9) $m(i \wedge j)=1$，则 $m(i)=m(j)=1$。

证明(1)：(i)对 $\mathscr{L}_N$，若 $\models i \longleftrightarrow j$，则据定义3.6，定理3.4，$RA(i)=RA(j)$，据定义3.12，$m^N(i)=m^N(j)$。(ii)对 $\mathscr{L}_\infty$，若 $\models i \longleftrightarrow j$，则据定理3.10(2)，易证存在一 n，使得对每一 $\mathscr{L}_{N^*}$ $(N^* \geqslant n)$，$\mathscr{L}_{N^*} \models i \longleftrightarrow j$。因此据(i)，$m^{N^*}(i)=m^{N^*}(j)$，再据极限定义，$m^\infty(i)=m^\infty(j)$。

证明(2)：若 $\models \neg i$，则 i 在 $\mathscr{L}_N$ 中逻辑假，因此据定义3.12(1)，$m^N(i)=0$，据极限定义，$m^\infty(i)=0$。

证明(4)：(i)对 $\mathscr{L}_N$，据定义3.11和3.12，$0 \leqslant m^N(i) \leqslant 1$。(ii)对 $\mathscr{L}_\infty$，据定义3.14，$0 \leqslant m^\infty(i) \leqslant 1$。

证明(6)：(i)对 $\mathscr{L}_N$，$RA(i \vee j)=RA(i) \cup RA(j)$，因此

$$
\begin{aligned}
m^N(i \vee j) &= \sum_{SD \in RA(i) \cup RA(j)} m^N(SD) = \sum_{SD \in RA(i)} m^N(SD) \\
&\quad + \sum_{SD \in RA(j)} m^N(SD) - \sum_{SD \in RA(i) \cap RA(j)} m^N(SD) \\
&= m^N(i)+m^N(j)-m^N(i \wedge j).
\end{aligned}
$$

(ii)对 $\mathscr{L}_\infty$，则据极限定义立证。

证明(8)：(i) 对 $\mathscr{L}_N, m(\neg i) = \sum_{SD \in RA(\neg i)} m(SD)$，据定理 3.1(5)，$= \sum_{SD \in VSD-RA(i)} m(SD) = \sum_{SD \in VSD} m(SD) - \sum_{SD \in RA(i)} m(SD) = 1 - m(i)$。(ii) 对 $\mathscr{L}_\infty$，据极限定义。

定理 3.13 对任意 $\mathscr{L}$ 若 (i) 对 $\mathscr{L}_N$，c 的变目中出现的句子是 $\mathscr{L}_N$ 的句子且第二变目的句子在 $\mathscr{L}_N$ 不能逻辑假。(ii) 对 $\mathscr{L}_\infty$，c 的变目中出现的句子是 $\mathscr{L}_\infty$ 的句子且 c 有值。(iii) 每一 c 表达式的值作为分母时是正的。则

(1) $0 \leqslant c(h,e) \leqslant 1$。

(2) 若 $\models e \to h$，则 $c(h,e) = 1$。

(3) 若 $\models \neg(e \wedge h)$，则 $c(h,e) = 0$。

(4) 若 $\models e_1 \leftrightarrow e_2$，则 $c(h,e_1) = c(h,e_2)$。

(5) 若 $\models h_1 \leftrightarrow h_2$，则 $c(h_1,e) = c(h_2,e)$。

(6) 若 $\models e \to \neg(h \wedge i)$，则

$$c(h \vee i, e) = c(h, e) + c(i,e).$$

(7)
$$c(h \wedge i, e) = c(h,e) \cdot c(i, e \wedge h) = c(i, e) \cdot c(h, e \wedge i).$$

(8) $c(\neg h, e) = 1 - c(h,e)$。

其中 (1),(2),(4),(6) 和 (7) 就是一般的概率演算用于描述概率函数的五条公理。

证明(1)：对 $\mathscr{L}_N$：因为 $m^N(e) \neq 0$，所以

$$c(h, e) = m^N(h \wedge e)/m^N(e).$$

因为 $\models h \wedge e \to e$，所以据定理 3.12(5)，$m^N(h \wedge e) \leqslant m^N(e)$，因此据定理 3.12(9)，$0 \leqslant \frac{m^N(h \wedge e)}{m^N(e)} \leqslant 1$。从而 $0 \leqslant c^N(h, e) \leqslant 1$。对 $\mathscr{L}_\infty$：据定义 3.15 即证。

证明(2)：若 $\models e \to h$，则 $\models h \wedge e \leftrightarrow e$。据定理 3.12

(5)，$m(h\wedge e)=m(e)$，所以 $c(h, e)=m(h\wedge e)/m(e)=m(e)/m(e)=1$。

证明(3)：若 $\models\neg(e\wedge h)$，则据定理 3.12 (2)，$m(e\wedge h)=0$。因此对 $\mathscr{L}_N$，$c^N(h,e)=m^N(h\wedge e)/m^N(e)=0$。对 $\mathscr{L}_\infty$，据定义 3.15，(3) 得证。

证明(4)：对 $\mathscr{L}_N$，因为据题设，有 $\models h\wedge e_1\leftrightarrow h\wedge e_2$。再据定理 3.12 (5)，$m^N(h\wedge e_1)=m^N(h\wedge e_2)$，所以有

$$
\begin{aligned}
c^N(h,e_1) &= m^N(h\wedge e_1)/m^N(e) \\
&= m^N(h\wedge e_2)/m^N(e) \\
&= c^N(h,e_2).
\end{aligned}
$$

(5) 的证明同上。

证明(6)：对$\mathscr{L}_N$，若$\models e\rightarrow\neg(h\wedge i)$，则$\models\neg((h\wedge e)\wedge(i\wedge e))$。据定理 3.12(2)，$m^N(h\wedge e\wedge i\wedge e)=0$。所以据定理 3.12(6)，我们有

$$
\begin{aligned}
c^N(h\vee i,e) &= m^N((h\vee i)\wedge e)/m^N(e) \\
&= m^N((h\wedge e)\vee(i\wedge e))/m^N(e) \\
&= (m^N(h\wedge e)+m^N(i\wedge e) \\
&\quad - m^N(h\wedge e\wedge i\wedge e))/m^N(e), \\
&= m^N(h\wedge e)/m^N(e) \\
&\quad + m^N(i\wedge e)/m^N(e) \\
&= c^N(h,e)+c^N(i,e).
\end{aligned}
$$

证明(7)：对 $\mathscr{L}_N$，假设 $m(e\wedge h)\neq 0$，

$$
\begin{aligned}
c^N(h\wedge i,e) &= m^N(h\wedge i\wedge e)/m^N(e) \\
&= \frac{m^N(h\wedge e)}{m^N(e)}\cdot\frac{m^N(i\wedge e\wedge h)}{m^N(h\wedge e)} \\
&= c^N(h,e)\cdot c^N(i,e\wedge h).
\end{aligned}
$$

若 $m^N(e\wedge h)=0$，则 $m^N(e\wedge h\wedge i)=0$，因此 $c^N(h\wedge i,e)=m^N(h\wedge i\wedge e)/m^N(e)=0$。而另一方面，$c^N(h,e)=0$，

因此仍有 $c^N(h\wedge i,e)=c^N(h,e)\cdot c^N(i,e\wedge h)$. 同理可证 $c(h\wedge i,e)=c(i,e)\cdot c(h,e\wedge i)$. 对 $\mathscr{L}_\infty$，根据定义 3.15，(7) 可得证.

证明 (8)：因为 $\models e\rightarrow\neg(h\wedge\neg h)$，所以据 (2), (6), (5) 及题设可得

$$c(h,e)+c(\neg h,e)=c(h\vee\neg h,e)$$
$$=c(\neg(h\wedge\neg h),e)=1.$$

从以上定理易得下列重要定理.

定理 3.14 若 $h_1,h_2,\cdots,h_n(n\geqslant 2)$ 相对 e 两两逻辑不相容，则

$$c(h_1\vee h_2\vee\cdots\vee h_n,\ e)=\sum_{p=1}^{n}c(h_p,\ e).$$

定理 3.15 $c(i,e)=c(h\wedge i,e)+c(\neg h\wedge i,e)$.

定理 3.16 (Bayes 定理) 令 $c(i,e)>0$，令 $h_1,h_2,\cdots,h_n(n\geqslant 2)$ 满足下列条件：(i)，$\models e\wedge i\rightarrow h_1\vee h_2\vee\cdots\vee h_n$，(ii) $e\wedge i\wedge h_1,e\wedge i\wedge h_2,\cdots,e\wedge i\wedge h_n$两两逻辑不相容。令 h 是 n 个 h 句中任一个，则

(1) $$c(h,e\wedge i)=c(i\wedge h,e)\Big/\sum_{p=1}^{n}c(i\wedge h_p,e)$$

$$=\frac{c(h,e)\cdot c(i,e\wedge h)}{\sum_{p=1}^{n}c(h_p,e)\cdot c(i,e\wedge h_p)}.$$

(2) 令 $c(h_p,e)$ 对每一 p (从 1 到 n) 有相同的值，则

$$c(h,e\wedge i)=\frac{c(i,e\wedge h)}{\sum_{p=1}^{n}c(i,e\wedge h_p)}.$$

显然，以上定义和定理适用性较差，除了一些特殊场合，

还不足以对 $\mathscr{L}$ 的任一句子对 h, e (e 非逻辑假)，确定 $c(h,e)$ 的值，因此满足定理 3.12 的 c 函数的个数是一连续统，其中有些 c 函数并无通常的归纳含义。例如对有些函数，相对正特例增多，确证度反而减少，而这显然不符合 Carnap 建立归纳逻辑（在 Carnap 看来概率逻辑就是归纳逻辑）的初衷，所以 Carnap 继续构造一特殊 c 函数（c^* 函数）来完成预期的目标。

定义 3.16 （1）对任意 $\mathscr{L}^\pi$ ($\mathscr{L}^\pi_N$ 或 $\mathscr{L}^\pi_\infty$)，$\mathscr{L}^\pi$ 的 Q 谓词表达式是 $P'_1 \wedge P'_2 \wedge \cdots \wedge P'_\pi$，其中每一 $P'_i (i \leqslant \pi)$ 是一元初始谓词 P_i 或 $\neg P_i$，我们用"Q_m"表示第 m 个（按词典顺序排列）Q 谓词，它表示的性质称为 Q 性质（显然对 $\mathscr{L}^\pi$，Q 谓词的数目 $K = 2^\pi$）。

（2）称 M 是一 Q 公式，若 M 是一 Q 谓词的原子公式（例 $Q_m(x)$ 是一 Q 公式）。

（3）称 i 是一 Q 句，若 i 是一 Q 谓词和个体常元的良构组合（例如，令 a 是 $\mathscr{L}^\pi$ 的任一个体常元，Q_m 是一 Q 谓词，则 $Q_m(a)$ 是一 Q 句，也称 a 有 Q_m 性质.）。

显然，对 $\mathscr{L}^\pi_N$，Q 句的总数 $\xi = K^N = (2^\pi)^N$。

例 3.1 设 P_1, P_2, P_3 是 $\mathscr{L}^3$ 的一元初始谓词，$\mathscr{L}^3$ 的 Q 谓词表如下：

Q 谓词表达式	Q 谓词
$P_1 \wedge P_2 \wedge P_3$	Q_1
$P_1 \wedge P_2 \neg \wedge P_3$	Q_2
$P_1 \wedge \neg P_2 \wedge P_3$	Q_3
$P_1 \wedge \neg P_2 \wedge \neg P_3$	Q_4
$\neg P_1 \wedge P_2 \wedge P_3$	Q_5
$\neg P_1 \wedge P_2 \wedge \neg P_3$	Q_6
$\neg P_1 \wedge \neg P_2 \wedge P_3$	Q_7
$\neg P_1 \wedge \neg P_2 \wedge \neg P_3$	Q_8

定理 3.17 对任意 $\mathscr{L}^{*}$，有

（1）由任意两不同的 Q 谓词和一个体常元构成的两 Q 句的合取是逻辑假的.

（2）具有同一个体常元的所有 Q 句的析取是逻辑真的.

（3）每一 Q 句是事实的.

（4）Q 谓词形成一划分.

（5）每一分子谓词(初始谓词和 $\neg$，$\vee$，$\wedge$ 的布尔组合）总可以用 $n(n \leqslant K)$ 个 Q 谓词表达式的析取来表达，从而每一分子句(即分子谓词表达式（用分子谓词来指称）和一个体常元的良构式）逻辑等价这 n 个 Q 句的析取(例如，对上述 $\mathscr{L}^{3}$, $P_1(a)$ 逻辑等价 $Q_1(a) \vee Q_2(a) \vee Q_3(a) \vee Q_4(a)$, $P_2(a) \wedge \neg P_3(a)$ 逻辑等价 $Q_2(a) \vee Q_6(a)$.).

（6）任意 n 个 Q 句的析取非逻辑假.

定义 3.17 令 $M(x_i)$ 是 $\mathscr{L}^{*}$ 的一公式，x_i 是它的唯一的自由变元，

（1）称 $M(x_i)$ 有逻辑宽度（简称宽度）ω，若（i）$M(x_i)$是逻辑空的(即 $(x_i)\neg M(x_i)$ 是逻辑真的)，且 $\omega=0$，或（ii）$M(x_i)$ 逻辑等价(称 $M_k(x_i)$ 逻辑等价 $M_j(x_i)$，若 $\models (x_i)(M_k(x_i) \longleftrightarrow M_j(x_i))$）含有 x_i 的Q公式且 $\omega=1$，或（iii）$M(x_i)$ 逻辑等价含有 x_i 的 ω 个不同Q公式的析取 $(\omega > 1)$.

（2）称 $M(x_i)$ 有相对逻辑宽度（简称相对宽度）q，若 $M(x_i)$ 有宽度 ω 且 $q=\omega/K$ $(K=2^{\pi})$.

定理 3.18 令 M 是 $\mathscr{L}^{*}$ 中具有宽度 ω，即有相对宽度 $q=\omega/K$ 的分子谓词表达式或分子谓词，则有

（1）M 是逻辑空的（因此由它和个体常元构成的分子句是逻辑假的）$\Longleftrightarrow \omega=0$，因而 $q=0$.

（2）M 是逻辑全称的(即若 M 为 $M(x_1, \cdots, x_n)$ 则

$\models (x_1)\cdots(x_n)\, M\ (x_1,\cdots,\ x_n)$，因此由它和个体常元构成的分子句是逻辑真的）$\Longleftrightarrow \omega = K$，因而 $q=1$.

(3) M 是事实的（因此由它和个体常元构成的分子句是事实的）$\Longleftrightarrow 0<\omega<K$，因而 $0<q<1$.

(4) $\neg M$ 有宽度 $K-\omega$，从而相对宽度为 $1-q$.

定义 3.18 $\mathscr{L}_N^\pi$ 的状态描述 SD 的Q形式定义为与 SD 逻辑等值的N个Q句的合取，其中每一Q句含有 $\mathscr{L}_N^\pi$ 的一个体常元，Q句的次序排列根据个体常元的下标的增长.

定义 3.19 $\mathscr{L}_N^\pi$ 的状态描述 SD 的Q数记为 $N_1,N_2,\cdots,N_K$，其中第m个Q数定义为在 SD 的Q形式中有 Q_m 性质的个体的数目.

因为共有K个Q性质，所以 $\mathscr{L}_N^\pi$ 的任一状态描述 SD 能确定它所描述的世界状态中有多少个体有 Q_1 性质，…，多少个体有 Q_K 性质. 显然 $\mathscr{L}_N^\pi$ 的状态描述 SD 的Q数有穷，它们的和就是 N. 显然对 $\mathscr{L}_N^\pi$ 的任一状态描述 SD，总有一些Q数等于 0.

定理 3.19 对任意 $\mathscr{L}_N^\pi$，令ζ是 $\mathscr{L}_N^\pi$ 的状态描述的数目，τ是 $\mathscr{L}_N^\pi$ 的结构描述的数目，则有

(1) $\zeta = K^N = (2^\pi)^N$. （因为 $\mathscr{L}_N^\pi$ 的原子句的数目是 πN，所以基本句的个数也为 πN，所以 $\zeta = 2^{\pi N} = K^N$）.

(2) $\tau = (N+K-1)!/N!(K-1)!$.

说明：$\mathscr{L}_N^\pi$ 的结构描述可以用由N个逗号和$K-1$条斜线组成的级数模式（serial pattern）来表示. 具有 Q_1 性质（用 Q_1 谓词来表示）的个体的个数由第一条斜线前的逗号来表示，具有 Q_2 性质的个体的个数由第一条斜线和第二条斜线之间的逗号来表示，由此类推，具有 Q_K 性质的个体的个数由第$K-1$条斜线后的逗号来表示（例如“,,,//,/,,,,/,,/,”指称具有 6 个性质的个体的个数分别为3,0,1,4,2,1），

所以 $\mathscr{L}_N^\pi$ 的结构描述的个数等于由N个逗号和$K-1$条斜线构成的模式的可能的个数,生成这些模式可以从 $N+K-1$ 个逗号的级数 (series) 开始,然后用斜线来置换其中的$K-1$个逗号组成的子类。这样的子类的数目是 $(N+K-1)!/N!(K-1)!$,从而可能的模式的个数也是 $(N+K-1)!/N!(K-1)!$。

定理 3.20 令 $\mathscr{L}_N^\pi$ 的任一状态描述 SD 的Q数为 $N_1, N_2, \cdots, N_K$, 又令 ST(SD) 是对应 SD 的结构描述, ξ 是 ST (SD) 的析取肢个数,也是 $\mathscr{L}_N^\pi$ 中与 SD 同构的状态描述的个数,则

$$\xi = N!/N_1!\cdots N_K!.$$

证明: 对 $\mathscr{L}_N^\pi$ 中的 $a_1, \cdots, a_N$ 和 Q_1 性质,$\cdots$,Q_K 性质(分别用 $Q_1, \cdots, Q_K$ 来表示),令具有 Q_1 性质的个体数为 $N_1, \cdots$, 具有 Q_K 性质的个体数为 N_K。显然, $N_1, \cdots, N_K$ 就是Q数,且 $N_1 + \cdots + N_K = N$。若给定 $a_1, \cdots, a_N$, 即对N个个体的 SD, 一一对应的置换共有 $N!$ 种。但是对 Q_1 谓词来说, $N_1!$ 种通过词典次序的重排就是它自身。由此类推,对 Q_K 谓词来说, $N_K!$ 种通过词典次序的重排就是它自身,所以 $N!$ 要除以 $N_1!\cdots N_K!$。

定义 3.20 称m是 $\mathscr{L}_N$ 的状态描述的对称测度函数,若m象定义 3.11 所定义,且 SD_i 同构于 SD_j (即 SD_j 与 SD_i 之间存在一 C对应,且 SD_i 按词典顺序重排),则

$$m(SD_i) = m(SD_j).$$

定义 3.20 实际是对传统无差别原则的一种定量的刻划, Carnap 认为它是无差别原则的有效内核,有关内容可参见[5], p.27。

我们可以用类似定义 3.12 和 3.13 的方法定义关于句子的对称测度函数和对称确证函数。

显然对称函数构成正则函数的一子类，所以关于后者的全部定理对前者也成立。下面我们讨论只对对称函数成立的定理。

定理 3.21 令 m 和 c 是对称函数，$i, h, e \in S$ ($\mathscr{L}$)；令 e 在 $\mathscr{L}$ 中非逻辑假，令 i', h', e' 分别是 i, h, e 的 **C** 对应，则下式成立：

(1) $m(i') = m(i)$.

(2) $c(h, e) = c(h', e')$.

定理 3.22 令 m 是 $\mathscr{L}_N^\pi$ 的句子的对称测度函数，且 c 是对应 m 的对称确证函数，令 SD 是 $\mathscr{L}_N^\pi$ 的任一状态描述，令 ξ 是 ST(SD) 的析取肢的数目，则有

(1) $m(\mathrm{ST}(\mathrm{SD})) = \xi \cdot m(\mathrm{SD})$.

(2) $c(\mathrm{SD}, \mathrm{ST}(\mathrm{SD})) = 1/\xi$.

证明 (1)：因为 ST(SD)的析取肢都是逻辑不相容的，所以据定理 3.13 就有上述结果。

证明(2)：因为 $\models \mathrm{SD} \rightarrow \mathrm{ST}(\mathrm{SD})$，所以

$$m(\mathrm{ST}(\mathrm{SD}) \wedge \mathrm{SD}) = m(\mathrm{SD}),$$

所以，据定义 3.13 和 (1)，有

$$c(\mathrm{SD}, \mathrm{ST}(\mathrm{SD})) = m(\mathrm{SD} \wedge \mathrm{ST}(\mathrm{SD}))/m(\mathrm{ST}(\mathrm{SD}))$$
$$= m(\mathrm{SD})/m(\mathrm{ST}(\mathrm{SD})) = 1/\xi.$$

接下来 Carnap 又从对称函数类中选取一个特定的函数 m^* 作为他的定量归纳逻辑的基础。

定义 3.21 称 m 是 m^* 函数，若 m 满足以下两个条件

(1) m 是对称测度函数。

(2) 若 ST 和 ST′ 是 $\mathscr{L}_N^\pi$ 的任意两结构描述，则

$$m(\mathrm{ST}) = m(\mathrm{ST}').$$

从定义 3.21 中我们看到，m^* 实际上是把概率 1，按 VSD ($\mathscr{L}_N^\pi$) 中同构类，即结构描述的个数 τ 等分地 ($1/\tau$) 赋于每

一同构类。然后据定义 3.20，再在每一同构类中把概率 $1/\tau$，按该类中状态描述的个数 ξ 等分地（$1/\tau\xi$）赋于每一状态描述。这样的赋值是比较自然的。

定理 3.23 对每一 $\mathscr{L}_N^\pi$，有且仅有一 m^* 函数。

证明：存在性显然，因为 m^* 是正则函数。现证唯一性。设 m^* 和 $m^{*\prime}$ 是两个满足定义 3.20 的函数，显然

$$\mathrm{dom}(m^*) = \mathrm{dom}(m^{*\prime}),$$

所以只须证，对任意 $i \in S(\mathscr{L}_N^\pi)$，

$$m^*(i) = m^{*\prime}(i).$$

令 $[\mathrm{SD}_j] = \{\mathrm{SD} \in \mathrm{RA}(i)\colon \mathrm{SD}$ 与 $\mathrm{SD}_j(\in \mathrm{RA}(i))$ 同构$\}$。因此存在一个 n，使得

$$\mathrm{RA}(i) = \bigcup_{j=1}^{n} [\mathrm{SD}_j],$$

即把 $\mathrm{RA}(i)$ 的 SD 划分成 n 个同构类，则根据定义 3.20 和 3.21，有

$$\begin{aligned}
m^*(i) &= \sum_{\mathrm{SD}\in \mathrm{RA}(i)} m^*(\mathrm{SD}) \\
&= \sum_{\mathrm{SD}\in[\mathrm{SD}_1]} m^*(\mathrm{SD}) + \cdots + \sum_{\mathrm{SD}\in[\mathrm{SD}_n]} m^*(\mathrm{SD}) \\
&= |[\mathrm{SD}_1]| \cdot m^*(\mathrm{SD}_1) + \cdots \\
&\quad + |[\mathrm{SD}_n]| \cdot m^*(\mathrm{SD}_n) \\
&= \frac{|[\mathrm{SD}_1]|}{\tau\xi_1} + \cdots + \frac{|[\mathrm{SD}_n]|}{\tau\xi_n},
\end{aligned}$$

其中 ξ_i 是 $\mathrm{VSD}(\mathscr{L}_N^\pi)$ 中与 SD_i 同构的状态描述的个数（$i = 1,\cdots,n$）。

同理，$m^{*\prime}(i) = \dfrac{|[\mathrm{SD}_1]|}{\tau\xi_1} + \cdots + \dfrac{|[\mathrm{SD}_n]|}{\tau\xi_n}$。

因此，$m^*(i) = m^{*\prime}(i)$。

定义 3.22 $\mathscr{L}_N^\pi$ 的 c^*, $\mathscr{L}_\infty^\pi$ 的 m^*, c^* 的定义分别类似定义 3.13,定义 3.14 和定义 3.15.

定理 3.24 对任一 $\mathrm{SD} \in \mathrm{VSD}(\mathscr{L}_N^\pi)$, 若 $N_1, N_2, \cdots, N_K$ 是 SD 的Q数,其中 $K = 2^\pi$, 则

$$m^*(\mathrm{SD}) = \frac{N_1! N_2! \cdots N_K! (K-1)!}{(N+K-1)!}.$$

证明:根据定义 3.20 和 3.21,定理 3.19 和 3.20,我们有

$$m^*(\mathrm{SD}) = 1/\tau\xi = \frac{N_1! \cdots N_K! (K-1)!}{(N+K-1)!}.$$

Carnap 的定量归纳逻辑最终选择 c^* 函数作为归纳确证度函数,因此他称他自己的定量归纳逻辑为 c^* 理论。

§2 归纳方法的一维连续统理论

Carnap 认为归纳推理在从物理学到历史学的所有经验科学领域里都起着非常重要的作用。这些领域科学探究的中心是假说或互相竞争的假说集。 假说可以是对单个事件(明天的天气、实验的结果、下届总统选举的结果)的预测,可以是一般趋势(癌症死亡率的下降、失业的增加)的预测,也可以是物理学或心理学或经济学的一条定律。在科学研究中,科学家经常做出关于假说(根据得到的证据)的判断,并且决定是否接受或拒斥它,这样的判断显然具有归纳本性。科学家意识到,在任何将来时刻,当发现新的证据,他将修改他对该假说的判断。新证据对该假说可能是有利的、不利的或不相干的。这里所谓有利的是指该假说得到比以前更强的确证。若这种确证增加的足够大,则原先被拒斥的假说现在可以接受。此外,科学家也可以对某个量的未知值进行定量估计。例如,根据得到的证据,对明天雨量进行估计。估计也是一种归纳

过程，因为我们在估计中不能保证该量的实际值接近估计值，更不能保证它等于估计值。

Carnap 要研究的归纳方法就是指归纳确证方法和归纳估计方法。这里只考察前者。

我们说一个人 X 有一种确证方法(归纳方法)，若他有能力以某种方式，至少在某种场合，根据证据 e 确证假说 h 的数值，把该数值看作 h 相对 e 的确证度。假设 X_1 使用 $c_1(h,e)$ 表示确证方法，X_2 使用 $c_2(h,e)$ 表示确证方法。若至少存在一句子对 h，e，使得 $c_1(h,e) \neq c_2(h,e)$，我们说这两种方法不相容。我们经常看到两种确证方法之间的差异，有时甚至会出现这样的情况，一种方法把 c 的极大值赋于给定的句子对 h，e，另一种方法却把 c 的极小值赋于同样的 h,e。例如，令 h 是无穷论域上的一全称事实句(例"所有的天鹅都是白的")，e 描述的是一有穷的正事例样本，于是一种方法(若后面将要讨论的直接规则)取 $c(h, e) = 1$，而另一种方法(即 c^*) 却取 $c(h,e) = 0$。

Carnap 的工作就是把可能的归纳方法用一定的方式排序为一系统。每一具体方法的本性决定它在该系统中的位置。反之，在该系统中也可用坐标或参数值描述某个方法的位置来唯一确定该方法。Carnap 就是要构造归纳方法的这样一个参数系统，他称其为 λ 系统。

Carnap 构造 λ 系统的语言是 $\mathscr{L}_N^\pi$，不特别考虑 $\mathscr{L}_\infty^\pi$，因为在涉及无穷论域的归纳方法中，c 的值是有穷系统序列 $\mathscr{L}_N^\pi$ 的值的极限(当 $N\to\infty$时)。

在构造 λ 系统之前，我们先研究可能的归纳方法。这将会导出一数学函数类，使得每一归纳方法正好为该类中一函数所刻划。一确证方法由一函数 $c(h,e)$ 来描述，但 c 不是数学函数，虽然它的值是数，但它的变目是句子。因此我们第

一个目标就是寻找另一种途径来刻划任意给定的确证方法，即用数学函数来代替 c 函数。

我们假设要包括在 λ 系统中的任意的确证方法可以用满足下列条件 3.1 至条件 3.10 的 c 函数来表示，其中条件 3.1 至条件 3.9 在此表述，条件 3.10 在后面表述。

条件 3.1 若 h 和 h' 逻辑等价，则 $c(h, e) = c(h', e)$。

条件 3.2 若 e 和 e' 逻辑等价，则 $c(h,e) = c(h, e')$。

条件 3.3 （一般乘法原则）$c(h \wedge h', e) = c(h, e) \cdot c(h', e \wedge h)$。

条件 3.4 （特殊加法原则）若根据 e，h 和 h' 逻辑不相容（即 $e \wedge h \wedge h'$ 逻辑假），则

$$c(h \vee h', e) = c(h, e) + c(h', e).$$

条件 3.5 $0 \leqslant c(h, e) \leqslant 1$。

显然条件 3.1—3.5 就是§1的定理 3.13 的 (5)，(4)，(7)，(6)，(1)。

为了表述条件 3.6—条件 3.9，我们先引入一些规定特殊形式的句子的记号。令 e_Q 是由 s 个不同的个体常元构成的 Q 句的合取，其中 s_1 个常元有 Q_1 性质，…，s_K 个有 Q_K 性质，因此 $s_1 + \cdots + s_K = s$。令 h_1 是 Q_1 句，其中的个体常元不在 e_Q 中出现，$h_2, \cdots, h_K$ 分别是同一常元和 $Q_2, \cdots, Q_K$ 谓词组成的 Q 句。这样，e_Q 可以看作是描述规模 (size) 为 s 的观察样本，$h_i (i \leqslant K)$ 可以看作是对未观察的个体的预测。令 e_1 是用 $\neg Q_1$ 去置换 e_Q 中不是 Q_1 的每一Q 谓词而得到的句子（例如“$Q_1(a_2) \wedge Q_3(a_5)$”被置换成“$Q_1(a_2) \wedge \neg Q_1(a_5)$”），令 $e_2, \cdots, e_K$ 也类似定义。这样任何形如 e_i 的句子描述了与 e_Q 相同的样本，但是用一种更少规定的方式，它只把 Q_i 性质赋于与 e_Q 相同的 s_i 个个体，对其余的 $s - s_i$ 个个体，只是说它们不是 Q_i，而不说明它

们的其它Q性质.

令M是有逻辑宽度$\omega(0 < \omega < K)$的事实分子谓词,因此M是ω个Q谓词的析取. 对M中的每一Q谓词,为了方便就说它属于M. 令e_M是用M去替换e_Q中的属于M的每一Q谓词(对于不属于M的Q谓词用$\neg M$去替换)而得到的句子. 这样,e_M是s_M个“M”的M句和$s - s_M$个“M”的M句的否定的合取,其中

$$s_M = \sum_{i \in M} s_i \text{(其中求和是对}M\text{中}Q\text{谓词的下标}i\text{)}. \quad (3.1)$$

这样e_M描述的是和e_Q相同的样本,但它只是说每一个体是否是M(当M是Q_i时e_i是e_M的特例). 令h_M是与h_1有相同个体常元的M句,因此,h_M逻辑等值ω个假说h_i(其中h_i中的Q谓词在M中)的析取,因为这些假说互相逻辑不相容,所以根据条件3.4,我们有

$$c(h_M, e) = \sum_{i \in M} c(h_i, e), \text{对任一 } e. \quad (3.2)$$

为了更具一般性,我们这里允许空样本的情况,即$s = 0$. 在这种情况下,e_Q(同样还有e_i和e_M)没有给出事实信息,所以它是逻辑真的句子,即是重言式“t”.

从描述一样本的证据到关于不属于该样本的个体的假说的归纳推理,也即对$c(h_1,e_1)$, $c(h_M,e_M)$, $\cdots$的确定,就是单称预测推理. Carnap 认为这是归纳逻辑中最基本的推理,别的类型的归纳推理在下列意义上都能归约为这种推理:若对$c(h_i,e_i)$的情况能给定一c函数的值,则所有别的值都能推导出来.

下面我们规定某些更具体的条件来刻划λ系统的c函数.

条件3.6 对给定的句子h_i和e_i,$c(h_i, e_i)$的值在所有h_i, e_i出现的系统$\mathscr{L}_N^\pi$中相同,且独立于N.

条件 3.7 c 相对个体常元是对称的.

条件 3.8 c 相对 Q 谓词是对称的,即若 h',e' 分别是从 h,e 中通过交换 (exchange) 两个 Q 谓词而得到的,则 $c(h,e)=c(h',e')$.

条件 3.9 对任意 M, $c(h_M,e_M)=c(h_M,e_Q)$.

从条件 3.9, 用 Q_i 代替 M, 对任意的 i, 就有

$$c(h_i,\ e_i)=c(h_i,e_Q). \tag{3.3}$$

从条件 3.9 和(3.2):

$$\text{对任意 } M, c(h_M,e_M)=\sum_{i\in M}c(h_i,e_Q)=\sum_{i\in M}c(h_i,\ e_i). \tag{3.4}$$

假设给定一确证方法(能运用于形如 e_1 和 h_1 那样的句子), 但这里我们不能根据一系列情况对 c 的值进行枚举计算, 因为可能的情况有无穷多. 因此必须给出一个一般的规则,构成 c 函数的定义. 这样在 $\mathscr{L}_N^\pi$ 中,对形如 h_1 和 e_1 的任意两句子, c 的值必须是这两个句子和该系统确定的某些量的函数. 现在让我们考虑可能作为该函数的变目的量. 首先, N和 π 是用来描述 $\mathscr{L}_N^\pi$ 的, 但是根据条件 3.6, N与 $c(h_1,\ e_1)$ 不相干. $\mathscr{L}_N^\pi$ 的Q谓词的个数K由 π 唯一确定 ($K=2^\pi$), 且反之亦然,因此我们可以把K看作是一变目,这样的K可以看作是逻辑的量.这里涉及的经验的量是由 e_1 描述的观察样本中的 s 和 s_1,因为 e_1 包含个体常元,它告诉我们 s 个个体属于该样本且其中哪几个属于 Q_1. 但由于条件 3.7 后一信息是独立的. 在 h_1 中出现的个体常元不同于 e_1 中出现的,但据条件 3.7, 它是哪个常元无关紧要. 因此, 对 h_1 的考虑不会引入任何其他的量. 这样,我们得知 $c(h_1,e_1)$ 的值由 K, s 和 s_1 唯一确定. 类似的结果对任何别的 Q_i 也

成立，因此

对 λ 系统的任意函数 c，存在一数学函数 G，使得对任意 Q_i 和任意具有 N,K,s 和 s_i 的 $h_i,\ e_i\in S(\mathscr{L}_N^{*})$，

$$c(h_i,e_i)=G(K,s,s_i). \qquad (3.5)$$

我们从 (3.5) 和 (3.4) 得到

$$\text{对任意 } M,\ c(h_M,e_M)=\sum_{i\in M}G(K,s,s_i). \qquad (3.6)$$

再据 (3.3)，得到

$$\text{对任意 } i,\ c(h_i,e_Q)=G(K,s,s_i). \qquad (3.7)$$

我们称 G 为上述给定的确证方法的特征函数。若给定各种确证方法(用 $c,c',\cdots$ 来表示)，则据 (3.5) 我们能确定它们的特征函数 $G,G'\cdots$。G 函数就是从整数三元组到实数闭区间 $[0,1]$ 的数学函数。

给定 K 个含有同一个体常元的句子 $h_i(i\leqslant K)$，令 h 是它们的析取，因为 h 逻辑真，所以相对任一证据，它的 c 值总为 1。此外，这些句子互不相容，因此据条件 3.4，有

$$\sum_{i=1}^{K}c(h_i,e_Q)=c(h,e_Q)=1.$$

因此，用"t"置换 e_Q (即 $s=s_i=0$)，得

$$\sum c(h_i,t)=1. \qquad (3.8)$$

据条件 3.8，

$$c(h_1,t)=c(h_2,t)=\cdots=c(h_K,t). \qquad (3.9)$$

因此，由上可得

$$\text{对 } \lambda \text{ 系统中任意 } c,\ Q_i,h_i,\ c(h_i,t)=1/K. \qquad (3.10)$$

再据 (3.4) 可得

$$\text{对任意 } c \text{ 和任意具有宽度为 } \omega \text{ 的 } M,\ c(h_M,t)=\omega/K. \qquad (3.11)$$

从(3.10)和(3.5)，对 $s=s_i=0$，我们有

$$G(K,0,0)=1/K, \text{对 } \lambda \text{ 系统中任意特征函数 } G. \quad (3.12)$$

若在我们后面的讨论中不能对 $G(K,0,0)$ 提供一个值，则我们有下列约定，即 G 函数的极限约定

$$\lim_{\delta\to 0} G(K,\delta,\delta)=G(K,0,0). \quad (3.13)$$

为了规定 λ 系统中这个给定的 G，根据(3.12)，用上述方式确定的值必须等于 $1/K$.

下面我们证明特征函数 $G(K,s,s_i)$ 完全刻划了它的确证方法. 这就意味，若给定一函数 G，我们能对任意 $\mathscr{L}_N^*$ 中任意句子对 h,e(e 非逻辑假)确定 c 的值.

令 $G(K,s,s_i)$ 是一给定的特征函数，我们要找到根据 G 如何确定函数 c 的值的方法. 显然，任意句子的 c 值可以归约为相对重言证据"t"的 c 值. 对后者的值，我们引入

$$m(h)=c(h,t). \quad (3.14)$$

这样，对每一 c 函数，存在由(3.14)定义的对应的测度函数，而任意句子的测度值又可以归约为状态描述的测度值，所以我们先考虑状态描述. 考虑具有 Q 形式的状态描述，即 $\mathscr{L}_N^*$ 中 N 个个体的 N 个 Q 句的合取 SD. SD 中组元的排序按下列形式：首先是 N_1 个 Q_1 句(其中 N_1是 Q_1 的 Q 数，它们之间的排序可以是任意的)，其次是 N_2 个 Q_2 句，由此类推，最后是 N_K个 Q_K句. 显然 $N_1+N_2+\cdots+N_K=N$. 令在这个序中的 Q 句是 $j_1,j_2,\cdots,j_N$，则 $\mathrm{SD}=j_1\wedge j_2\wedge\cdots\wedge j_N$. 令 SD_n 是 SD 的前 n 个 j 句的合取：$\mathrm{SD}_n=j_1\wedge j_2\wedge\cdots\wedge j_n$，因此，当 $n\geqslant 2$ 时，$\mathrm{SD}_n=\mathrm{SD}_{n-1}\wedge j_n$. 据条件 3.3，我们有

$$c(\mathrm{SD}_n,t)=c(\mathrm{SD}_{n-1},t)\cdot c(j_n,\mathrm{SD}_{n-1}). \quad (3.15)$$

对 n 重复上述过程，得

$$m(\mathrm{SD}) = c(\mathrm{SD}, t) = c(j_1, t)c(j_2, \mathrm{SD}_1)\cdots c(j_N, \mathrm{SD}_{N-1}),$$

其中

$$\mathrm{SD} = \mathrm{SD}_N, \mathrm{SD}_1 = j_1. \tag{3.16}$$

考虑 SD 中出现的一特定谓词“Q_i”. SD 中包含 N_i 个具有 Q_i 性质的句子. 令 m_i 是 SD 中先于第一个 Q_i 句的合取肢的个数. 因此,若 $i = 1$, 则 $m_i = 0$, 若 $i > 1$,则

$$m_i = \sum_{p=1}^{i-1} N_p.$$

这 N_i 个 Q_i 句是 $j_{m_i+1}, j_{m_i+2}, \cdots, j_{m_i+N_i}$.

考虑 (3.16) 右边乘积中的项, 我们首先假设“Q_i”不是 SD 中出现的第一个 Q 谓词, 因此 $m_i > 0$, 这样 (3.16) 右边确证“Q_i”的第一个项是 $c(j_{m_i+1}, \mathrm{SD}_{m_i})$, 其中 SD_{m_i} 是 m_i 个不含“Q_i”的 Q 句的合取. 这里 c 的两变目具有形如 h_i 和 e_Q 的句子,其中 $s = m_i$, $s_i = 0$. 因此,根据 (3.6),第一个项是 $G(K, m_i, 0)$. 另一方面,若“Q_i”是 SD 中第一个 Q 谓词, 因而 $m_i = 0$, 则这样的第一个项是 $c(j_1, t)$, 据 (3.10), 它总为 $1/K$, 这里 $s = s_i = 0$. 这样的第一个项可表述为 $G(K, 0, 0)$, 这和当 $m_i = 0$时, $G(K, m_i, 0)$ 的结果是相一致的. 第二个项, 不论 $m_i = 0$ 还是 $m_i > 0$, 都是 $c(j_{m_i+2}, \mathrm{SD}_{m_i+1})$, 这里情况类似上述情况,只是 $s = m_i + 1$, $s_i = 1$, 因此,第二个项是 $G(K, m_i + 1, 1)$. 用这种方法继续下去, 在每一步, s 和 s_i 都增长 1. 对最后一个有“Q_i”的项, $s = m_i + N_i - 1$, $s_i = N_i - 1$, 因此, 最后一个项是 $G(K, m_i + N_i - 1, N_i - 1)$. 这样在 (3.16) 中具有“$Q_i$”的项的乘积是 $\prod_{p=1}^{N_i} G(K, m_i + p - 1, p - 1)$, 同理, 对别的 Q 谓词,也有类似的结果成立,因此,据 (3.16),对任意状态描述 SD (其 Q 数为 $N_i(i \leqslant K)$),我们有

$$m(\mathrm{SD}) = \prod_{i} \prod_{p=1}^{N_i} G(K, m_i + p - 1, p - 1), \quad (3.17)$$

其中当 $N_i > 0$,第一个乘积符 $\prod_{i}$ 的 i 等于 $i=1,2,\cdots,K$,在整个乘积中第一项总为 $1/K$.

这样,G 对所有的状态描述确定了测度值,从而任何别的句子的测度值都能确定,因为据条件 3.4 (特殊加法原则),可以证明,对任意 $j \in S(\mathscr{L}_N^\pi)$,

$$\begin{cases} m(j) = 0, \text{ 若 } j \text{ 逻辑假}, \\ m(j) = \sum\limits_{\mathrm{SD} \in \mathrm{RA}(j)} m(\mathrm{SD}), \text{ 否则}. \end{cases} \quad (3.18)$$

因此所有的 c 值都可以归约为

$$c(h,e) = m(e \wedge h)/m(e), \text{ 对任意 } h,e \in \mathrm{S}(\mathscr{L}_N^\pi), \quad m(e) \neq 0. \quad (3.19)$$

这样,对任意 $h, e \in \mathrm{S}(\mathscr{L}_N^\pi)$,只要 e 非逻辑假,特征函数能确定它的 c 值.

因此,可以在下列意义上证明,特征函数 G 唯一确定它的 c 函数. 令 c 是满足条件 3.1—3.9 的 c 函数,且对任意 $h,e \in \mathrm{S}(\mathscr{L}_N^\pi)$ (其中 e 非逻辑假),c 值存在. 令对应的特征函数 $G(K,s,s_i)$ 由 (3.5) 确定,若我们对任意 h, e 运用在 (3.17),(3.18),(3.19) 所描述的过程,则 (3.19) 就会产生给定的函数 c 对 h,e 的值 $c(h,e)$. 尤其是考虑形如 h_i, e_i 的句子对,则 (3.19) 就会根据 (3.5) 来确定 $G(K, s, s_i)$ 的值 $c(h_i,e_i)$. 根据本章第一节建立的 c 函数理论,这是很容易得到的. 事实上,这个结果意味,任意 G 函数至多确定一满足上述条件的 c 函数(显然,并不是每一任意选出的函数 $G(K, s,s_i)$ 都能确定满足上述条件的 c 函数)

这样,我们解决了我们问题的第一部分,对每一确证方

法，存在一完全刻划此方法的数学函数 G. 现在我们进到第二个,也是更困难的部分：为了完全刻划任意函数 G(从而完全刻划对应的归纳方法)，我们希望找到一参数集,使得这样的参数个数不仅有穷,而且还很少,否则就不适合我们预想的目的：即通过研究已知和未知的归纳方法的特征(由它们的参数完全确定这种特征)来对这些方法进行定量的描述。

考虑G的变目和值的适域 (range)，因为 π 是一正整数,所以 $K=2^{\pi}$ (Q 谓词的数目)是 2 的正乘幂. s 是由 e_i 描述的样本中的个体数,因此 s 是小于N的非负整数(通常是正的,但也允许 $s=0$). s_i 是小于或等于 s 的非负整数，而G的值就是 c 的值，因此应该是闭区间 $[0,1]$ 中的实数. 假设一科学家X已做了某一观察,在他的证据 e_M 中,他描述了一给定样本中的个体. 他想知道根据 e_M，关于性质M的单称预测 h_M 的 c 值. 让我们看看在这种情况下可能影响 c 值的两特定的因素. 我们把 s_M/s (在 e_M 描述的观察样本中M出现的频率)作为第一个因素,这个因素是经验因素，它一般为大家接受，并且被看作是归纳推理的基本特性之一. 第二个要考虑的因素是M的相对宽度,即 ω/K. 这个因素是逻辑因素，独立于任何观察事实. X 在给定的语言系统 $\mathscr{L}_{\pi}^{N}$ 中进行逻辑分析就能确定 ω/K(给定的"M"的定义),同时X能把"M"变形为某些Q谓词的析取,析取肢的个数就是 ω，而 $K=2^{\pi}$. 显然，经验因素 s_M/s 和逻辑因素 ω/K 都影响 c 的值.

现在我们考虑G的取值区间. 显然 s_M/s 和 ω/K 这两个因素有特殊的重要性,因为 $c(h_M,e_M)$ 的值必须始终在这两个值之间或等于它们中的一个. Carnap 认为这个条件为所有迄今已有的归纳方法所满足，作为恰当性的必要条件似乎也是合乎情理的,因此 λ 系统的最后一条件是:

条件 3.10 对任意具有 K,s,s_i 的句子 h_i,e_i，或 $s_i/s \leqslant c(h_i,e_i) \leqslant 1/K$，或 $1/K \leqslant c(h_i,e_i) \leqslant s_i/s$.

这样，λ 系统包括了所有满足条件 3.1—3.10 的 c 函数. 在下列意义上，Carnap 认为他的 λ 系统的元素是完全归纳方法. 这就是,若在任意 $\mathscr{L}_N^\pi$ 中对任一句子对 h,e (e 非逻辑假)，c 有值，则语言系统 $\mathscr{L}_N^\pi$ 的一确证方法以及它的函数 c 称为完全的.

对任意分别有 K,ω,s,s_M 的M和句子 h_M,e_M，从条件 3.10,(3.1) 和 (3.4)，能得到对M的一般条件

$$s_M/s \leqslant c(h_M,e_M) \leqslant \omega/K, \text{ 或 } \omega/K \leqslant c(h_M,e_M) \leqslant s_M/s. \quad (3.20)$$

因为 $c(h_i, e_i)$ 给出 $G(K,s,s_i)$ 的值，所以对 λ 系统的任意特征函数 G 和任意 K,s,s_i，从条件 3.10 可得

$$s_i/s \leqslant G(K,s,s_i) \leqslant 1/K, \text{ 或 } 1/K \leqslant G(K,s,s_i) \leqslant s_i/s. \quad (3.21)$$

我们打算把 $G(K,s,s_i)$ 的值表示为值 s_i/s 和 $1/K$ 的加权平均,根据约定把经验因子 s_i/s 的权标准化，然后用逻辑因素 $1/K$ 的权确定 $G(K,s,s_i)$ 的值. 为此选择观察样本的个体数 s 作为 s_i/s 的标准权. 这个选择无需理论上的正当性 (justification)，因为它不涉及任何断定,它只是一个约定. 这个选择实践上的正当性是基于如下事实：它将导致特别简单的参数系统.

把经验因素的权标准化之后,准备赋于逻辑因素 $1/K$ 的权就具有函数G的特性. 对一给定的G，这种权无需总是同一个,而是可以随情况而变,这种变化可以依赖值 K,s,s_i，因此我们把它当作这三个变目的函数 $\lambda(K,s,s_i)$. 这样G的值是 s_i/s 和 $1/K$ 的加权平均，其中第一个权是 s，第二个是 $\lambda(K,s,s_i)$，因此，

$$G(K,s,s_i) = \frac{s_i + \lambda(K,s,s_i)/K}{s + \lambda(K,s,s_i)}, \quad (3.22)$$

即

$$\lambda(K,s,s_i)=\frac{s\cdot G(K,s,s_i)-s_i}{1/K-G(K,s,s_i)}. \tag{3.23}$$

这样，若任意归纳方法用函数 c 的形式给定，那么我们可以先确定它的特征函数 G，然后根据(3.23)确定它的函数 λ. 另一方面，若给定函数 λ，则可以确定函数 c. 这样，G 函数和对应的 λ 函数互为确定，因此后者也可看作是对应的归纳方法的特征函数，就象 G 函数那样. 但这两个函数有很大的区别，通过去掉三个变目中任一个来简化函数 G 的一般形式是不可能的，因为确定 $G(K,s,s_i)$ 的值的 $c(h_i,e_i)$ 的值一般说依赖这三个变目.

另一方面，使用函数 λ 会导致极大的简化，因为我们可以去掉三个变目中的一些变目. 若对任意我们熟悉的归纳方法，我们能根据前面的过程构造它的函数 G，从而据(3.23)构造它的函数 λ，那么我们发现后者根本没有包含变目 s 和 s_i（这在后面讨论具体的归纳推理时能看到）.

一给定的函数 λ 不依赖 s 和 s_i，这并不意味由这个 λ 函数刻划的归纳方法，在确定 $c(h_i,e_i)$ 时不能考虑 s 和 s_i，而是象(3.22)说明的那样，s 和 s_i 在确定 $c(h_i,e_i)$ 的函数 G 中出现，即使它不在 λ 函数中出现. 尤其 s 不出现在 λ 函数中是由于我们选择 s 作为经验因素的标准权. s 的不出现只说明我们选择这个权极好地符合习惯上的归纳方法的共同特征.

当我们允许 λ 系统包括所有满足条件 3.1—3.10 的归纳方法时，为了简单起见，我们先研究其 λ 函数独立于 s 和 s_i 的归纳方法.

条件 3.11 在后面的讨论中，假定，若给定一 $\mathscr{L}^{\pi}_{N}$，则对任意 s，s_i，h_i 和 e_i，$(s\cdot c(h_i,e_i)-s_i)/(1/K-c(h_i,e_i))$

是一常数。

由于条件 3.11，函数 λ 只剩一个变目 K，因此我们现在可以用 $\lambda(K)$ 来代替 $\lambda(K,s,s_i)$，这样 (3.22) 等就成为

$$\left.\begin{matrix} G(K,s,s_i) \\ c(h_i,e_i) \end{matrix}\right\} = \frac{s_i + \lambda(K)/K}{s + \lambda(K)}. \quad \begin{matrix}(3.24)\\(3.25)\end{matrix}$$

因此，据 (3.4) 和 (3.1)，

$$c(h_M,e_M) = \frac{s_M + (\omega/K)\lambda(K)}{s + \lambda(K)}. \tag{3.26}$$

(3.26) 中的经验因素和逻辑因素分别是 s_M/s 和 ω/K，权同于前面。

现在我们证明，λ 函数完全刻划了它的归纳方法。给定一函数 $\lambda(K)$，为了简单起见，我们用 λ 来代替 $\lambda(K)$，则据 (3.17) 和 (3.24)，我们有

$$m(\mathrm{SD}) = \prod_i \prod_{p=1}^{N} \frac{p-1+\lambda/K}{m_i+p-1+\lambda}, \tag{3.27}$$

其中当 $m_i=0$，$p=1$ 时，商总取 $1/K$。

因此

$$m(\mathrm{SD}) = \prod_i \mathrm{A}_i \Big/ \prod_i \mathrm{B}_i, \tag{3.28}$$

其中 A_i 和 B_i 规定如下：

$$\mathrm{A}_i = \prod_{p=1}^{N_i} (\lambda/K + p - 1), \tag{3.29}$$

$$\mathrm{B}_i = \prod_{p=1}^{N_i} (\lambda + m_i + p - 1). \tag{3.30}$$

从 (3.29) 得

$$\mathrm{A}_i = \lambda/K \cdot (\lambda/K+1)(\lambda/K+2)\cdots(\lambda/K+N_i-1). \tag{3.31a}$$

令 i_1 和 i_2 是 i 的两个值,使得 $i_1 < i_2$, $N_{i_1} > 0$, $N_{i_2} > 0$,若 $i_1 < i < i_2$, 则 $N_i = 0$. 对 $i > 0$, $m_i = \sum_{p=1}^{i-1} N_p$, $m_{i_2} = m_{i_1} + N_{i_1}$, 根据 (3.30), 对任意 i, B_i 是 N_i 个上升因子的乘积,其中每一因子比前面的因子大 1. B_{i_1} 中最大的因子是 $\lambda + m_{i_1} + N_{i_1} - 1$, B_{i_2} 中最小的因子是 $\lambda + m_{i_1} + N_{i_1}$, 因此后者比前者大 1,所以 $\prod_i B_i$ 是 $\Sigma N_i = N$ 个从 λ 到 $\lambda + N - 1$ 的上升因子的乘积. 因此从 (3.28) 和 (3.31a),

$$m(\mathrm{SD}) = \prod_i \frac{((\lambda/K)(\lambda/K+1)\cdots(\lambda/K+N_i-1))}{\lambda(\lambda+1)\cdots(\lambda+N-1)}. \tag{3.31b}$$

这样就为归纳方法建立了一 λ 系统。

现在我们考察 λ 函数的变化，以及由此产生的两类不同的归纳方法,并且考察其中的几个重要的方法.

第一类归纳方法,即 λ 独立于 K. Carnap 认为几乎所有的归纳方法的 λ 函数不仅独立于 K，而且有一常数值，例外的只是他提出的确证函数 c^*. 这里考虑的 λ 不仅可以是所有的有穷正实数（对应于区间 $[\omega/K, s_M/s]$ 中的内点), 而且可以是 0 或∞(对应于这个区间的端点)，为此我们先给出以下对任意函数 $f(\lambda)$ 的极限约定:

$$\begin{cases} f(\lambda) = \lim\limits_{\lambda\to 0} f(\lambda), & \lambda = 0 (f(\lambda)\text{的定义没有指派值给 } f(\lambda)), \\ f(\lambda) = \lim\limits_{\lambda\to\infty} f(\lambda), & \lambda = \infty. \end{cases} \tag{3.32}$$

我们引入一些由给定的 λ 刻划的函数的记号. G_λ 是由 λ 根据 (3.24) 和(3.32)确定的 G 函数，c_λ 是由 G_λ 刻划的 c 函数，m_λ 是对应 c_λ 的测度函数,于是 (3.25) 和 (3.26) 就有下列较简单的形式

$$c_\lambda(h_i, e_i) = \frac{s_i + \lambda/K}{s + \lambda}, \tag{3.33}$$

$$c_\lambda(h_M, e_M) = \frac{s_M + (\omega/K)\lambda}{s + \lambda}. \tag{3.34}$$

在证据是重言的,即样本是空的情况下, 从 (3.10) 和 (3.11), 得

$$m_\lambda(h_i) = c_\lambda(h_i, t) = 1/K, \tag{3.35}$$

$$m_\lambda(h_M) = c_\lambda(h_M, t) = \omega/K. \tag{3.36}$$

Carnap 在此分别详细地讨论了几种第一类的归纳 方 法 (参见[4]的 §12—§13). $\lambda = 2$ 的归纳方法是一种修正了的 Laplace 规则, $\lambda = 1$ 的归纳方法赋予逻辑因素和个体的观察以相同的权. 而当 $\lambda = \infty$, 根据极限约定 (3.32), 我们从 (3.24)、(3.25) 和 (3.33) 得出

$$G_\infty(K, s, s_i) = \lim_{\lambda \to \infty} \frac{s_i + \lambda/K}{s + \lambda} = 1/K. \tag{3.37}$$

这样, G_∞ 等于逻辑因素, 完全独立于 s 和 s_i, 即经验因素被忽略。

据 (3.37) 或直接从 (3.25),(3.26) 和 (3.33),

$$c_\infty(h_i, e_i) = 1/K, \tag{3.38}$$

$$c_\infty(h_M, e_M) = \omega/K. \tag{3.39}$$

(3.39) 中的 c_∞ 等于M的相对权, 与 e_M 描述的观察样本所包含的信息不相干. 从 (3.38) 和 (3.3),

$$c_\infty(h_i, e_Q) = 1/K. \tag{3.40}$$

令 SD 是 $\mathscr{L}_N^\pi$ 的任一状态描述, (3.16) 右边的每一乘积因子对 c_∞ 都有 (3.40) 那样的形式,因此等于 $1/K$, 所以

$$m_\infty(\mathrm{SD}) = 1/K^N. \tag{3.41}$$

而 K^N 恰是 $\mathscr{L}_N^\pi$ 的状态描述的数目,所以 m_∞ 把等概分配给 $\mathscr{L}_N^\pi$ 的每一状态描述. 这种归纳方法可以称为随机归纳 方法, 但 Carnap 认为这种归纳方法是不可接受的。

当 $\lambda = 0$ 时，假设 $s > 0$，我们有

$$c_0(h_i, e_i) = s_i/s, \tag{3.42}$$

$$c_0(h_M, e_M) = s_M/s. \tag{3.43}$$

我们称 (3.43) 为直接规则 (straight rule)，它表明单称预测推理 c_0 等于观察频率。当 $\lambda = 0$ 时，(3.35) 和 (3.36) 也成立，因此

$$m_0(h_i) = c_0(h_i, t) = 1/K, \tag{3.44}$$

$$m_0(h_M) = c_0(h_M, t) = \omega/K. \tag{3.45}$$

令 SD 是 $\mathscr{L}_N^\pi$ 的具有 Q 数 $N_i (i \leqslant K)$ 的状态描述、运用极限约定于 (3.31b)，就有

$$m_0(\text{SD}) = \lim_{\lambda \to 0} \frac{\prod_i (\lambda/K(\lambda/K+1)\cdots(\lambda/K+N_i-1))}{\lambda(\lambda+1)\cdots(\lambda+N-1)}, \tag{3.46}$$

其中乘积的取值范围是 $I = \{i : N_i > 0\}$。 令 $p = |I|$，这样，上述乘积包含 p 个乘积因子，因此

$$\begin{aligned} & m_0(\text{SD}) \\ &= \lim_{\lambda \to 0} \left(\frac{(\lambda/K)^p}{\lambda} \cdot \frac{\prod_i (\lambda/K+1)\cdots(\lambda/K+N_i-1)}{(\lambda+1)\cdots(\lambda+N-1)} \right) \\ &= \lim_{\lambda \to 0} \frac{(\lambda/K)^p}{\lambda} \cdot \lim_{\lambda \to 0} \frac{\prod_i (\lambda/K+1)\cdots(\lambda/K+N_i-1)}{(\lambda+1)\cdots(\lambda+N-1)} \\ &= \lim_{\lambda \to 0} \left(\frac{(\lambda/K)^p}{\lambda} \right) \cdot \frac{\prod_i (N_i-1)!}{(N-1)!}. \end{aligned} \tag{3.47}$$

我们分两种情况讨论。情况 1：令 $p = 1$，这时 SD 中每一个体都有相同的 Q 谓词，因为它们相对 $\mathscr{L}_N^\pi$ 中可表达的性质是完全相同的，所以我们称这种情况下的 SD 是齐次 (ho-

mogeneous) 状态描述. 这时只有一个正 N_i，因此它等于 N，所以 (3.47) 中右边最后一个商变成 $(N-1)!/(N-1)!$，而极限号下的商变成 $1/K$，因此，若 SD 是一齐次状态描述，则

$$m_0(\mathrm{SD}) = 1/K. \tag{3.48}$$

情况 2: $p>0$，(3.47) 极限号下的商变成 λ^{p-1}/K^p，因此极限为 0，而最后一个商总是有穷的，所以，若 SD 是一非齐次的状态描述，则

$$m_0(\mathrm{SD}) = 0. \tag{3.49}$$

据该结果和 (3.18)，对 $\mathscr{L}_N^\pi$ 的许多事实句 i，$m_0(i)=0$，所以我们一般不能用 $m_0(e\wedge h)/m_0(e)$ 来确定 $c_0(h,e)$，因此不得不直接用极限约定于 c_0：

$$c_0(h,e) = \lim_{\lambda\to 0} c_\lambda(h,e) = \lim_{\lambda\to 0}\frac{m_\lambda(h\wedge e)}{m_\lambda(e)}. \tag{3.50}$$

对每一非逻辑假的 e 和每一 $\lambda>0$，根据(3.31b)，$m_\lambda(e)>0$，因此 (3.50) 中的商有一确定的值. 可以证明当 $\lambda\to 0$ 时，这些值收敛于一极限. 因此，对任意 $\mathscr{L}_N^\pi$ 的每一句子对 h，e(e 非逻辑假)，由 (3.50) 给出 $c_0(h,e)$ 的确定值. 对于 h_M 和 e_M，当 $s=0$ 时，(3.50) 就是 (3.45)；对于 h_M 和 e_M，当 $s>0$ 时，(3.50) 就是 (3.43).

我们知道，满足条件 3.1—3.5 且对每一状态描述有正值的测度函数以及依赖这种测度函数的 c 函数都称为正则的. Carnap 认为只有正则的 c 函数才表示恰当的归纳推理，根据 (3.31b) 可证，对任意有穷正数 λ，m_λ 是正则的. 因而 c_λ 是正则的. 根据 (3.41)，m_∞ 是正则的，因而 c_∞ 是正则的. 但从 (3.49) 可见，m_0 不是正则的，因此我们称 m_0 和 c_0 是拟正则 (quasi-regular) 函数，这是指它们自身不是正则的，但它们能表示为正则函数的极限而言的.

我们现在来看第二类归纳方法，即 λ 是 K 的函数的归纳方法，先看 $\lambda(K)=bK$（其中 b 是一常项系数)的一类较简单的方法．据 (3.31b)，我们有

$$m(\mathrm{SD})=\frac{\prod_i (b(b+1)\cdots(b+N_i-1))}{bK(bK+1)\cdots(bK+N-1)}. \tag{3.51}$$

从 (3.25) 和 (3.26)，

$$c(h_i,e_i)=\frac{s_i+b}{s+bK}, \tag{3.52}$$

$$c(h_M,e_M)=\frac{s_M+b\omega}{s+bK}. \tag{3.53}$$

如果 b 是正整数，(3.51) 就成为

$$m(\mathrm{SD})=\frac{(bK-1)!}{(bK+N-1)!}\prod_i\frac{(b+N_i-1)}{(b-1)!}. \tag{3.54}$$

我们考虑 (3.54) 的一重要特例

1. $b=1$，即 $\lambda(K)=K$，这时的归纳方法就是 c^* 方法：

$$m^*(\mathrm{SD})=\frac{(K-1)!}{(N+K-1)!}\prod_i N_i!. \tag{3.55}$$

$$c^*(h_i,e_i)=\frac{s_i+1}{s+K},$$

$$c^*(h_M,e_M)=\frac{s_M+\omega}{s+K}.$$

这样 c^* 表示第二类归纳方法中最简单的一种．

值得提一下，每一与 SD 同构（同构的状态描述有相同的 Q 数 $N_i(i\leqslant K)$）状态描述有相同的 m^* 值． 对应的结构描述 ST(SD) 是与 SD 同构的状态描述的析取．因此，据条件 3.4，$m^*(\mathrm{ST(SD)})$ 等于 $m^*(\mathrm{SD})$ 乘以与 SD 同构的状态描述的个数$\left(\text{即 } N_i\big/\prod_i N_i!\right)$．由此，我们有

$$m^*(\mathrm{ST}(\mathrm{SD})) = \frac{(K-1)!N!}{(N+K-1)!}. \qquad (3.56)$$

因为这个结果独立于Q数 N_i，所以每一结构描述有相同的 m^* 值。

2. $b = 2$，因此 $\lambda(K) = 2K$，据 (3.54)，

$$m(\mathrm{SD}) = \frac{(2K-1)!}{(N+2K-1)!} \prod_i (N_i+1)!. \qquad (3.57)$$

从(3.52)，(3.53)，可得

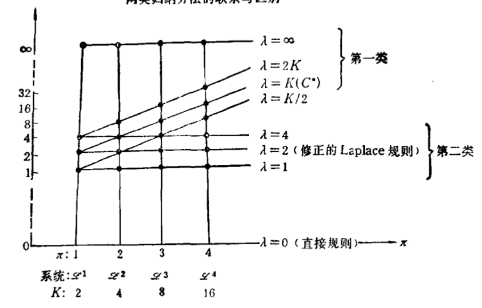
$$c(h_i, e_i) = \frac{s_i + 2}{s + 2K}, \qquad (3.58)$$

$$c(h_M, e_M) = \frac{s_M + 2\omega}{s + 2K}. \qquad (3.59)$$

当然人们可以考虑第二类归纳方法中一个更广泛的子类，包括K的非线性函数。Carnap 把它限制于线性函数的简单式是因为他认为目前对第二类归纳方法的研究还没有任何实际的理由要超出这一范围。

那么这两类方法之间有什么差异（对第二类方法我们只研究了 $\lambda(K) = bK$ 的情况）？在第一类方法中，参数λ的值本身是归纳方法的特性，与语言系统（模型论意义上）无关。而对第二类方法，λ是相对语言系统的。Carnap 本人认为这两类方法的差异不那么基本和重要。他选择几种代表性的归纳法列在一张示意图上表示它们的联系与区别。

λ系统用函数或数λ描述了无穷多的归纳方法。于是就产生了以下问题：对一个人X来说，若他根据自己的观察结果想要确定确证度，那么他应该选择何种有实效的方法。这基本上不是一理论问题，因为对该问题的回答在于X所做的实际决策。决策不能判断其真假，只能看作多少是恰当的，即对给定的目的是否适当。当然，选择的恰当性依赖关于各种归



纳法性质的理论上的结果，因而理论上的结果也可能影响他的决策，但决策本身仍是一实践问题. 首先 X 必须决定选择哪一类归纳方法，正象我们所见，第一类归纳方法在下述方面有其优越性：在计算 $c(h,e)$ 时，只有那些在 h 和 e 中出现的初始性质才须考虑.

假设 X 由于这样或那样的理由选择了一种第一类方法，那么他必须选择任一实数值，或∞，0 来作为参数 λ 的值. 他知道当 $\lambda=\infty$ 时的归纳方法的性质，例如 c_∞ 独立于 e. 假设他由于实践上的评价认为该方法不可用，假设他还拒绝使用 λ 非常大(虽然 λ 有穷)的那些方法，因为它们也有同样不令人满意的功用，虽然程度上小一些. 同样他也可能不用那些 λ 接近 0 的归纳方法. 他可能会感到小的 λ 值似乎能导致 c 的更恰当的值，整数会导致更简单的公式，特别是当 λ 是 2 的乘幂时，因此他可以首先考虑 $\lambda=2$，即修正了的 Laplace 方法. 或考虑 $\lambda=1$，再考虑 $\lambda=4$，$\lambda=8$，$\lambda=16$，…，也可以考虑 $\lambda=1/2$，$\lambda=1/4$，….

假设X宁可选择第二类归纳方法。这里只考虑 $\lambda(K)=bK$。若他把 b 看作是正整数，则除了 $\lambda=0$ 或 $\lambda=\infty$，该方法要比第一类方法简单得多。他首先可以考虑 $b=1$，因而 $\lambda(K)=K$，即函数 c^*。函数 c^* 的定义有极大的简洁性，它比别的要考虑的 c 函数更简单。其次他可以把 b 看作 2,3，也能考虑 3 以后的几个整数，但不能很多。因为若 b 的值比较大，且初始性质较多的系统，λ 的值将相当大，这样得到的 c 的值的恰当性将会十分可疑。例如，既使 $b=6,\pi=8$，则 $K=2^8=256,\lambda=1536$。这可能损害其恰当性，因此，第二类方法中似乎只有一小部分才有优越性。

假设X选择了某个归纳方法，在某段时间内把该方法用于出现的归纳问题。若他不满意该方法的结果，他随时可以摒弃它，继而选择对他更具吸引力的方法。他可以考虑归纳方法的性能，它提供的值，它与后面的实验结果的关系。例如预测的真值频率（truth-frequency）和误差估计，进而考虑使用该方法的经济性（这由所要求的计算的简单性来测度）。还可以考虑美学上的特性：像定义和规则在逻辑上的简洁性。λ 系统使我们容易找到另一种方法，因为它在一个尺度中提供了系统的排序方式。若X感到他目前使用的方法与逻辑因素相比，不能给经验因素以足够的权，则他可以根据是小一点还是小很多来选择具有较小 λ 的归纳方法。另一方面，若他希望给逻辑因素以更多的影响，给经验因素更少的影响，则他可以在 λ 尺度上向上移动他的刻度。

在结束本章时，我们对 Carnap 的工作作一简要的评价。Carnap 的归纳逻辑理论在现代归纳逻辑史上占极为重要的地位。他第一次用精妙的技巧以及逻辑语义学的方式构造了博大的归纳逻辑的理论。但是 Carnap 的工作也有重大缺陷，主要表现在：① 从连续统多个确证函数中挑选归纳确证

函数 c^* 的方法(指他在 §1 的工作)是特设的，人为性很大. ②更严重的是,对 c^*，一全称事实句在无穷个体域上相对有穷证据的归纳确证度总为 0 。这是相当致命的.尽管 Carnap 认为,归纳逻辑主要应关注单称预测推理而不是全称推理(就象在§2的工作),试图回避矛盾，但终究说明这样的归纳逻辑意义不大. 因为归纳逻辑有一个很重要的目的：即合理地解释、刻划科学方法论中证据对假说的确证度. 而在科学实践中,特别是在理论色彩较浓的自然科学的实践中,人们主要关注的是全称形式的假说。

由于 Carnap 的工作的成就和缺陷都非常突出，使得不少人不断地在修改和完善 Carnap 的理论,下面两章,我们分别从不同的侧面介绍两种对 Carnap 归纳逻辑理论的完善和扩展。

第四章　归纳方法的多维连续统理论

Carnap 的 λ 系统可以看作是一个一维连续统．前面已提及这个系统有重大缺陷．为了克服这个缺陷，J. Hintikka 和 I. Niiniluoto 提出了能很好地避免这个缺陷的多维连续统理论，发展了 Carnap 思想．在本章第一节中，我们介绍和我们 Hintikka 在[6]中提出的二维 α-λ 系统．在第二节中讨论介绍和讨论 Niiniluoto 在[7]中提出的 K 维系统．

§1　Hintikka 的二维 α-λ 连续统理论

J. Hintikka 认为 Carnap 的一维连续统理论之所以造成以下困难：在无穷论域上全称事实句相对有穷证据的归纳概率（c^*）为 0，是因为 Carnap 分配概率的方式有问题．从第三章我们已知 Carnap 分配概率给状态描述（它们最终决定句子对的概率）分两个层次：首先把概率 1 均分给结构描述，然后再在结构描述内部按状态描述的数目再均分已分配给结构描述的概率．因为结构描述的数目依赖语言系统给定的个体数目，所以当论域无穷时，势必要造成零概率的情况．Hintikka 认为，为了避免发生这类困难，就要改变分配概率的方式使得分配概率只依赖初始谓词的情况．

和 Carnap 一样，Hintikka 也从有 m 个初始一元谓词的语言出发（注意：这里的 x 表示空位，下同）：

$$P_1(x), P_2(x), \cdots, P_m(x). \tag{4.1}$$

如 Carnap 在第三章的构造，我们可以建立 $K=2^m$ 个 Q 谓

词：$Q_1, Q_2, \cdots, Q_K$. 用Q谓词，构造一类特殊的句子，Hintikka 称为构元（constituent）。一个构元陈述了，有且仅有若干个Q谓词被例证（称一Q谓词被例证，若论域中至少有一个体具有该Q谓词描述的性质）。在给出构元的形式定义之前，我们举例来说明构元的性质。

考虑只有两个一元初始谓词的语言：$B(x)$，$R(x)$. 为了便于理解，我们用 $B(x)$ 表示"x 是黑的"，$R(x)$ 表示"x 是一只乌鸦"，则我们有以下Q谓词：

$Q_1 = B(x) \wedge R(x)$（x 是一黑乌鸦），

$Q_2 = B(x) \wedge \neg R(x)$（$x$ 是一黑的非乌鸦），

$Q_3 = \neg B(x) \wedge R(x)$（$x$ 是一非黑的乌鸦），

$Q_4 = \neg B(x) \wedge \neg R(x)$（$x$ 是一非黑非乌鸦）.

通过断定哪些Q谓词正好被例证，我们可以从这些Q谓词中构造不同的构元。例如，给定一个黑乌鸦和一个非黑非乌鸦（例如一块白石头）的论域，则下列构元在该论域中真：

$$\exists x Q_1(x) \wedge \exists x Q_4(x) \wedge (x)(Q_1(x) \vee Q_4(x)).$$

这个构元可以记作C_2，它表示有且仅有 Q_1 和 Q_4 被例证。由此可见，从K个Q谓词中，选择不同的Q谓词我们可以构造不同的构元。一般我们用 C_w 表示恰有w个Q谓词 $Q_{i_1}, \cdots, Q_{i_w}$ 被例证的任意构元：

$$C_w \triangleq \bigwedge_{j=1}^{w} \exists x Q_{i_j}(x) \wedge (x)\left(\bigvee_{j=1}^{w} Q_{i_j}(x)\right).$$

Hintikka 用下列 Bayes 定理作为出发点来建立他的 α-λ 系统：

$$P(C_w, e_n) = \frac{P(C_w) \cdot P(e_n, C_w)}{\sum_{i=0}^{K-c} \binom{K-c}{i} P(C_{c+1}) P(e_n, C_{c+1})}, \quad (4.2)$$

其中P是满足第三章§2的条件3.1—3.5的概率测度，e_n表

述了有且仅有 n 个个体例证了下列 c 个 Q 谓词

$$Q_{i_1}(x), Q_{i_2}(x), \cdots, Q_{i_c}(x). \tag{4.3}$$

假设 e_n 中例证 $Q_{i_j}(x)$ 的个体数为 n_j，则

$$\sum_{j=1}^{c} n_j = n.$$

从(4.2)已知，只要我们能确定 $P(C_w)$ 和 $P(e_n, C_w)$ 的不同值，则(4.2)的值也随之确定，因此我们先来考虑先验概率 $P(C_w)$.

在 Carnap 的 λ 系统中，相对无穷论域，当 $w=K$时，$P(C_w)=1$，而当 $w\neq K$时，$P(C_w)=0$，因此所有的全称事实句的概率（无论是先验的还是后验的）都为 0. 为了避免这个结论，我们必须把不同的非零先验概率 $P(C_w)$ 分配给不同的构元. 显然，这样的分配应该只依赖 w. 为了与 Carnap 的方法作比较，我们先回顾一下 Carnap 的 λ 系统如何确定 $P(C_w)$. 给定一构元 C_w，与 C_w 相容的 α 个个体（它们例证 C_w 中被例证的 Q 谓词）的概率是

$$\frac{(w\lambda/K)}{\lambda}\cdot\frac{1+(w\lambda/K)}{1+\lambda}\cdot\cdots\cdot\frac{\alpha-1+(w\lambda/K)}{\alpha-1+\lambda}. \tag{4.4}$$

下面我们用 $\pi(\alpha, w\lambda/K)$ 和 $\pi(\alpha, \lambda)$ 分别指称(4.4)中分子的乘积和分母的乘积.

为了把(4.4)扩展为变目是 α, w, λ, K 的函数，很自然的一种方式是把(4.4)等同于

$$\frac{\Gamma(\alpha+w\lambda/K)\cdot\Gamma(\lambda)}{\Gamma(\alpha+\lambda)\cdot\Gamma(w\lambda/K)}, \tag{4.5}$$

其中 $\Gamma(x)$ 是通常的 γ 函数 (gamma-function).

当 α 很大时，在 α 个个体的论域中，(4.4)接近概率 $P(C_w)$. 当 α 较小时，这个 $P(C_w)$ 是

$$\sum_{i=0}^{w}(-1)^{i}\binom{w}{i}\frac{\pi(\alpha,(w-i)\lambda/K)}{\pi(\alpha,\lambda)}. \qquad (4.6)$$

显然，Hintikka 在建立自己的 α-λ 系统时，并不把 $P(C_w)$ 等同(4.6)，而是使它与(4.4)(或(4.5))成比例，即

$$P(C_w)=\pi\left(\alpha,\ \frac{w\lambda}{K}\right)\Big/\sum_{i=0}^{K}\binom{K}{i}\pi(\alpha,i\lambda/K), \qquad (4.7)$$

这里的 λ 原则上依赖 K，因此(4.7)可写作

$$P(C_w)=\pi\left(\alpha,\ \frac{w\cdot\lambda(K)}{K}\right)\Big/\sum_{i=0}^{K}\binom{K}{i}\pi\left(\alpha,\ \frac{i\lambda(K)}{K}\right). \qquad (4.7)^{*}$$

这就是在无穷论域上，Hintikka 对 C_w 的概率赋值。

那么(4.7)中的参数 α 究竟起什么作用呢？显然为了使 C_w（其中 $w<K$）真，论域中的个体正好例证了 w 个 Q 谓词。当 $P(C_w)$（其中 $w<K$）的概率越高，则在我们论域中的个体的分布就越不平均，离先验对称性要求就越远，这时的 α 就越小。这在一定意义上表明，α 是归纳概括句(事实句)中先验考虑的强度指标。因此，这里的 α 与 Carnap 的参数 λ 类似，因为后者也是归纳过程中先验(逻辑)因素的权的指标。λ 越大，单称预测的值越接近由于对称考虑所导致的先验值。同样，α 越大，$P(C_w)$（其中 $w<K$）越接近它的先验值 0．两个参数的区别只在于它们适合不同的归纳推理：α 适合归纳概括推理，而 λ 只适合单称预测推理。

现在我们来考虑(4.2)的另一概率值：似然度（degree of likelihood）$P(e_n, C_w)$。这相当于在固定了构元的先验概率后，考虑如何在所有的状态描述中把这个先验概率再分配给使该构元为真的状态描述。显然我们不能直接使用 Carnap 的特征函数

$$(n_i + \lambda/K)/(n + \lambda), \qquad (4.8)$$

因为这会导致与我们已分配给构元的先验概率不一致，但对(4.8)稍加修改就能满足我们的要求。因此我们用下列特征函数

$$(n_j + \lambda/w)/(n + \lambda), \qquad (4.9)$$

其中w就是在构元C_w中出现的Q谓词的数目。根据(4.9)，在无穷论域的情况下，我们得到下列$P(e_n, C_w)$:

$$\frac{\lambda/w}{\lambda} \cdot \frac{1+\lambda/w}{1+\lambda} \cdot \cdots \cdot \frac{n_1 - 1 + \lambda/w}{n_1 - 1 + \lambda} \cdot \frac{\lambda/w}{n_1 + \lambda}$$

$$\cdot \frac{1+\lambda/w}{n_1 + 1 + \lambda} \cdot \cdots \cdot \frac{n_2 - 1 + \lambda/w}{n_1 + n_2 - 1 + \lambda}$$

$$\cdot \frac{\lambda/w}{n_1 + n_2 + \lambda} \cdot \cdots = \left(\prod_{j=1}^{c} \pi(n_j, \lambda/w)\right) \Big/ \pi(n, \lambda). \qquad (4.10)$$

根据(4.7)*和(4.10)，我们能计算出构元C_w的后验概率（确证度）$P(C_w, e_n)$，和先验概率$P(e_n)$。也能计算出一未知个体具有Q谓词Q_{i_1}的后验概率，记之为$P(h, e_n)$，它相当于 Carnap 的单称预测推理

$$P(C_w, e_n) = \left[\left(\pi\left(\alpha, \frac{w \cdot \lambda(K)}{K}\right) \Big/ \pi(n, \lambda(w))\right) \cdot \prod_{j=1}^{c} \pi\left(n_j, \frac{\lambda(w)}{w}\right)\right] \Big/ \left[\sum_{i=0}^{K-c}\left(\binom{K-c}{i} \cdot \left(\pi\left(\alpha, \frac{(c+i)\cdot\lambda(K)}{K}\right) \Big/ \pi(n, \lambda(c+i))\right) \cdot \prod_{j=1}^{c} \pi\left(n_j, \frac{\lambda(c+i)}{c+i}\right)\right)\right], \qquad (4.11)$$

$$P(e_n) = \sum_{i=0}^{K-c}\left(\binom{K-c}{i} P(C_{c+i}) \cdot P(e_n, C_{c+i})\right)$$

$$= \left[\sum_{i=0}^{K-c}\left(\binom{K-c}{i}\pi\left(\alpha,\ \frac{(c+i)\cdot\lambda(K)}{K}\right)\right.\right.$$

$$\left.\left.\cdot\prod_{j=1}^{c}\pi\left(n_j,\ \frac{\lambda(c+i)}{c+i}\right)\Big/\pi(n,\lambda(c+i))\right)\right]\Big/$$

$$\sum_{i=0}^{K}\binom{K}{i}\pi\left(\alpha,\ \frac{i\cdot\lambda(K)}{K}\right). \tag{4.12}$$

而 $P(h, e_n)$ 就是 $P(h\wedge e_n)/P(e_n)$，其中 $P(h\wedge e_n)$ 类似(4.12)，除了 n_1 被 n_1+1 置换，n 被 $n+1$ 置换(这里当然预设 $n_1>0$)外。在(4.11)和(4.12)中，我们很自然假设 λ 不仅是 w 的函数，此外它还是 K 或 i 的函数。若假设 λ 是独立于 w 和 i 的常数，则我们有下列较简单的表达式

$$\mathrm{A}_i = \binom{K-c}{i}\pi(\alpha, (c+i)\lambda/K),$$

$$\mathrm{B}_i = \prod_{j=1}^{c}\pi(n_j, \lambda/(c+i)),$$

$$P(C_w, e_n) = \left[\pi(\alpha,\ w\lambda/K)\cdot\prod_{j=1}^{c}\pi(n_j, \lambda/w)\right]\Big/$$

$$\left[\sum_{i=0}^{K-c}(\mathrm{A}_i\cdot\mathrm{B}_i)\right], \tag{4.13}$$

$$P(e_n) = \sum_{i=0}^{K-c}\mathrm{A}_i\cdot\mathrm{B}_i\Big/\sum_{i=0}^{K}\left(\binom{K}{i}\pi(\alpha, i\lambda/K)\cdot\pi(n,\lambda)\right), \tag{4.14}$$

$$P(h,\ e_n) = \left[\sum_{i=0}^{K-c}\left(\mathrm{A}_i\cdot\prod_{j=1}^{c}\pi(n_j, \lambda/(c+i))\right.\right.$$

$$\left.\left.\cdot(n_1+\lambda/(c+i))\right)\right]\Big/\sum_{i=0}^{K-c}(\mathrm{A}_i\cdot\mathrm{B}_i)(n+\lambda). \tag{4.15}$$

(4.10)—(4.15) 刻划了 Hintikka 的 α-λ 系统在无穷论域上的特征。现在我们来讨论包含在 α-λ 系统中的几个子系统。先考虑 $\alpha\to\infty$的情况。为了清楚起见，我们用 $\pi(\alpha,\lambda)$ 把(4.15)的分子和分母分开．当 $\alpha\to\infty$时，

$$\frac{\pi(\alpha,(c+i)\lambda/K)}{\pi(\alpha,\lambda)}=\frac{(c+i)\lambda/K}{\lambda}\cdot\frac{1+(c+i)\lambda/K}{1+\lambda}\cdot\cdots\cdot\frac{\alpha-1+(c+i)\lambda/K}{\alpha-1+\lambda},\tag{4.16}$$

当 α 相对其它参数变得非常大时，下列(4.17)就近似(4.16)：

$$\frac{\Gamma(\lambda)}{\Gamma((c+i)\lambda/K)}\alpha^{((c+i)\lambda/K)-\lambda}.\tag{4.17}$$

当 $\alpha\to\infty$时，若 $c+i=K$，则(4.17)$\to 1$；若 $c+i\neq K$，则(4.17)$\to 0$。因为(4.15)的各和式中只有一项不等于0，所以(4.15)可以归结为

$$(n_1+(\lambda/K))/(n+\lambda).\tag{4.18}$$

这正是 Carnap 的 λ 系统的特征函数，因此 Carnap 的一维 λ 系统已作为特例包含在 Hintikka 的二维 α-λ 系统中了．例如当 $\alpha\to\infty$且 $\lambda=K$时，就可以从 α-λ 系统中得到 c^* 函数。

α-λ 系统的另一个有趣的特例是当 $\lambda\to 0$ 时的结果。根据 λ 的含义，这相当于根据完全后验(经验)的证据来判定单称归纳预测。在这种情况下，(4.15)的特征函数就有下列形式：

$$n_1/n.\tag{4.19}$$

这就是“直接规则”(这里我们假设 λ 不依赖 w)，请参见第三章§2的(3.42)和(3.43)。这里还须注意，当 α 有穷时，

(4.19)与α无关.这时,对于(已被e_n列举)在该论域中被例证的Q谓词的构元C_c, $P(C_c, e_n)=1$;对所有其它的构元C_w, $P(C_w, e_n)=0$.

若我们令$\alpha=0$,据(4.5),所有的先验概率$P(C_w)$与同一个常数成比例,因此都等于$1/2^K$. 事实上在这种情况下,(4.13)成为

$$\prod_{i=1}^{c}\pi(n_i,\lambda/w)\Big/\sum_{i=0}^{K-c}\binom{K-c}{i}B_i, \tag{4.20}$$

或更一般地,(4.13)成为

$$\left[\sum_{i=0}^{K-c}\binom{K-c}{i}\frac{(\Gamma(\lambda/w))^c}{(\Gamma(\lambda/(c+i)))^c}\cdot\prod_{j=1}^{c}\frac{\Gamma(n_j+(\lambda/(c+i)))}{\Gamma(n_j+(\lambda/w))}\right]^{-1}, \tag{4.20*}$$

且(4.15)成为

$$\left[\sum_{i=0}^{K-c}\left(\binom{K-c}{i}\prod_{j=1}^{c}(\Gamma(n_j+(\lambda/(c+i)))\right.\right.$$
$$\left.\left./\Gamma(\lambda/(c+i)))(n_1+(\lambda/(c+i)))\right)\right]$$
$$\Big/\left[\sum_{i=0}^{K-c}\left(\binom{K-c}{i}\prod_{j=1}^{c}\Gamma(n_j+(\lambda/(c+i)))\right.\right.$$
$$\left.\left./\Gamma(\lambda/(c+i))\right)(n+\lambda)\right]. \tag{4.21}$$

在此我们预设λ是常数. 当$w>c$且$n\to\infty$时,(4.20)*$\to 0$,当$w=c$且$n\to\infty$时,(4.20)*$\to 1$.

现在我们考虑$\lambda\to\infty$,即单称归纳预测推理只根据先验基础判定. 为了看清这一点,我们把(4.20)*近似写作

$$\left(\prod_{j=1}^{c}n_j\right)^{((1/(c+i))-1/w)\lambda}, \tag{4.22}$$

当 $\lambda\to\infty$ 时，(4.7)* 趋于

$$(w/K)^{\alpha}\Big/\sum_{i=0}^{K}\binom{K}{i}(i/K)^{\alpha}. \qquad (4.23)$$

用同样的方式，我们看到当 $\lambda\to\infty$ 时，(4.10) 趋于 $(1/w)^{n}$. 把这些值代入 (4.13) 且进行简化，我们就有下列 C_w 的确证度

$$P(C_w)=\left[\sum_{i=0}^{K-c}\binom{K-c}{i}\left(\frac{w}{c+i}\right)^{n-\alpha}\right]^{-1}. \qquad (4.24)$$

这是 Hintikka 在[8] 构造的 Jerusalem 系统关于构元的确证度的直接概括. 为了说明 (4.24) 中 α 的作用，我们考虑只有两个初始谓词 $R(x)$，$B(x)$ 的语言，若我们假设除了$Q_3=R(x)\wedge\neg B(x)$，其它所有的Q 谓词都被例证，则 $m=2$，$K=4$，$c=3$. 令 $C_3=\exists xQ_1(x)\wedge\exists xQ_2(x)\wedge\exists xQ_4(x)\wedge(x)(Q_1(x)\vee Q_2(x)\vee Q_4(x))$，据 (4.24)，得

$$P(C_3)=\left(1+\left(\frac{3}{4}\right)^{n-\alpha}\right)^{-1}. \qquad (4.25)$$

在这种情况下，α 就是我们打赌(公平地)剩下的乌鸦是黑之前必须观察的个体数.

前面我们考虑的所有特例都是通过给出 α 和 λ 的适当数值得到的. 但 λ 的值可以依赖 K，因此我们令 λ 依赖 w，即 $\lambda=\lambda(w)$ 来得到另一些系统. 在此我们考虑这样的系统中最简单的一个：$\lambda=w$. 在一定意义上，这是对 Carnap 的 c^* 系统 ($\lambda=K$) 的概括. 在这种情况下，从 (4.11) 得到下列简化形式

$$P(C_w,e_n)=\frac{(\alpha+w-1)!}{(n+w-1)!}\Big/\sum_{i=0}^{K-c}\binom{K-c}{i}\frac{(\alpha+c+i-1)!}{(n+c+i-1)!}. \qquad (4.26)$$

这是对 Hintikka 在 [9] 提出的"联合系统"的概括。当 $\alpha=0$ 时,这里的系统就是 [9] 中的"联合系统"。用同样的方式,我们有

$$P(h,e_n)=\frac{n_1+1}{n+c}\frac{\sum_{i=0}^{K-c}\binom{K-c}{i}\frac{(\alpha+c+i-1)!(n+c)!}{(n+c+i)!}}{\sum_{i=0}^{K-c}\binom{K-c}{i}\frac{(\alpha+c+i-1)!(n+c-1)!}{(n+c+i-1)!}}. \tag{4.27}$$

下面的图表说明上面我们讨论的各种子系统与 Hintikka 的 α-λ 系统的关系。

系　统	参　数	
	α	λ
Carnap 的 λ 系统	∞	λ
c^* 系统	∞	K
直接规则	α	0
Jerusalem 系统[8]	0	∞
Jerusalem 系统的概括	α	∞
依赖(4.20),(4.21)的系统	0	λ
"联合系统"[9]	0	w
"联合系统"的概括	α	w

表中右边两栏中出现的 α 或 λ 表示在该系统中它们是可以取任何值的参数。

§2　Niiniluoto 的 K 维连续统理论

K 维连续统理论最早是 J. Hintikka 和 I. Niiniluoto 在 [10] 中提出的。后来 I. Niiniluoto 在 [7] 对他和 Hintikka 在 [10] 中提出的理论进行概括和阐述,在此我们主要介

绍和讨论 [7] 中的工作。

考虑论域 U，通常 U 是无穷的。假设 U 可以划分成 K 个互不相交又联合穷举的 Q 类：$Q_1, Q_2, \cdots, Q_K$。显然 $Q_1, Q_2, \cdots, Q_K$，其中 $2^m = K$，可以看作是对"Q 谓词"的解释（如 Carnap 那样）。令 e_n 是从论域中（无置换）抽取的规模为 n（即 n 个元素）的样本，n_i 是 e_n 中属于类 Q_i 的个体数 $(i = 1, \cdots, K)$，因此 $n_i \geq 0$ 且

$$\sum_{i=1}^{K} n_i = n.$$

若对应论域 U 的个体 $u_1, u_2, \cdots$，语言 $\mathscr{L}$ 有个体常元符号 $a_1, a_2, \cdots$，则 e_n 可以用 $\mathscr{L}$ 的一特称句来表述。这样的句子也用"e_n"来指称。令 $h_{n+1} = Q_i(a_{n+1})$（注意：这里的 e_n 和 h_{n+1} 与本章 §1 的 e_n 和 h 本质上是一样的），P 是 $\mathscr{L}$ 的概率测度，则对 $\mathscr{L}$ 的所有句子对 $\langle h, e\rangle$，e 为非逻辑假，概率值 $P(h, e)$ 满足下列公理。

公理 4.1 概率公理。

公理 4.2 对有穷证据 e，若 $\vdash e \to h$，则 $P(h, e) = 1$。

公理 4.3 $P(h, e)$ 相对个体常元是对称的。

公理 4.4 $P(h, e)$ 相对 Q 谓词是对称的。

公理 4.5 （c 原则） $P(h_{n+1}, e_n)$ 是 n_i, n, c 和 K 的函数，但不是 n_j（其中 $j \neq i$）的函数，其中 $c = |\{i : n_i > 0\}|$。

公理 4.1—4.5 就是 Niiniluoto K 维系统的 5 条公理，其中公理 4.1—4.4 相当于第三章 §2 中 Carnap 的 λ 系统的条件 3.1—3.5 和条件 3.7—3.8。公理 4.5 是对 Carnap 的下列 λ 原则的概括：

$P(h_{n+1}, e_n)$ 是 n_i, n 和 K 的函数，但不是 n_j（其中 $j \neq i$）的函数。 (4.28)

从公理 4.1—4.4 和 (4.28) 可以得到下列基本结果：若

公理 4.1—4.4 和 (4.28) 成立，则 $P(h_{n+1}; e_n)$ 的特征函数可以表示为下列形式：

$$f(n_i, n) = \frac{n_i + \lambda/K}{n + \lambda}, \qquad (4.29)$$

其中 λ 是正实值参数，由方程

$$\lambda = \frac{K \cdot f(0,1)}{1 - K \cdot f(0,1)}$$

确定(参见第三章 § 2 的 (3.24) 和 (3.25))。

这里还要注意 c 原则弱于 (4.28). 为了理解 c 原则，我们先来看它的某些直接推论. 对固定的 K, $P(h_{n+1}, e_n)$ 的特征函数现在可以表示为三元函数 $f(n_i, n, c)$. 从公理 4.3 直接推出

若 $n' > 0$, $n'' > 0$, $K \geqslant c \geqslant 2$, $n - n' - n'' \geqslant c-2$, 且 $n = n' + n''$ (当 $c = 2$ 时)，则

$$\begin{aligned} & f(n', n, c) \cdot f(n'', n+1, c) \\ = & f(n'', n, c) \cdot f(n', n+1, c). \end{aligned} \qquad (4.30)$$

特别是当 $c = K$时，(4.30) 成立. 现在可以用下列结果重新证明 (4.29) 成立.

若 $n' \geqslant 1$, 且 $n - n' \geqslant K - 1$, 则

$$f(n', n, K) = \frac{n' + \lambda/K}{n + \lambda}, \qquad (4.31)$$

其中 $\lambda > -K$ 是用下列方程确定的常数

$$\lambda = \frac{K \cdot f(1, K+1, K)}{1 - K \cdot f(1, K+1, K)} - K.$$

当 n' 或 n'' 等于 0 时，(4.30) 不成立，因为在这种情况下，e_n 包括所有被例证的 Q 谓词. 我们可以从公理 4.3 得到下列结果来代替 (4.30).

若 $n' \geqslant 1$, $K \geqslant c \geqslant 1$, $n - n' \geqslant c - 1$, 且当 $c = 1$

时 $n' = n$，则

$$f(n',n,c)f(0,n+1,c)$$
$$= f(0,n,c)f(n',n+1,c+1). \qquad (4.32)$$

在 (4.32) 的假设条件下，重复使用 (4.32)，我们有

$$f(n',n,c) = \mu(n,c)f(n',n+K-c,K), \qquad (4.33)$$

其中

$$\mu(n,c) = \prod_{j=0}^{K-c-1} f(0,n+j,c+j)/f(0,$$

对 $c < K$，$\mu(n,K) = 1, n+1+j, c+j)$.

若使用记号

$$\alpha(n,c) = \mu(n,c)/(n+K-c+\lambda), \qquad (4.34)$$

则从 (4.31) 和 (4.32) 得

若 $n' \geqslant 1, n - n \geqslant c-1$，且当 $c=1$ 时 $n = n'$，则

$$f(n',n,c) = \frac{\mu(n,c)(n'+\lambda/K)}{n+K-c+\lambda}$$
$$= \alpha(n,c)(n'+\lambda/K). \qquad (4.35)$$

对 $n \geqslant c$，定义

$$h(n,c) = (K-c)f(0,n,c),$$
$$g(n,c) = 1 - h(n,c)$$
$$= 1-(K-c)f(0,n,c). \qquad (4.36)$$

函数 h 和 g 的直观意义是显然的。一个体（不在 e_n 中出现的个体）是属于一新种类（即它属于一个新 Q 类）的概率等于 $h(n,c)$，$g(n,c)$ 是下一个体属于 e_n 中已被例证的种类之一的概率。若 $n_1,\cdots,n_K$ 中的非零数是 $n_{i_1},\cdots,n_{i_c}$，则

$$g(n,c) = \alpha(n,c)(n+c\lambda/K)$$
$$= \mu(n,c)\cdot(n+c\lambda/K)/(n+K-c+\lambda), \qquad (4.37)$$
$$h(n,c) = 1-\alpha(n,c)(n+c\lambda/K), \qquad (4.38)$$
$$f(0,n,c) = \frac{1}{K-c}(1-\alpha(n,c)(n+c\lambda/K)). \qquad (4.39)$$

若 (4.35) 的假设成立，则有

$$g(n,c)=(n+c\lambda/K)\cdot f(n',n,c)/(n'+\lambda/K), \quad (4.40)$$

$$f(n',n,c)=(n'+\lambda/K)(1-(K-c)\cdot f(0,n,c))/(n+c\lambda/K). \quad (4.41)$$

特别是当 $n=c$ 时,

$$g(c,c)=c\cdot\alpha(c,c)(1+\lambda/K)=c\cdot\mu(c,c)/K$$
$$=1-(K-c)f(0,c,c)=c\cdot f(1,c,c). \quad (4.42)$$

把(4.35)和(4.39)代入(4.32),就得到 $\alpha(n,c)$ 的差方程

$$\alpha(n+1,c+1)(n+c\lambda/K-1/\alpha(n,c))$$
$$=\alpha(n+1,c)(n+1+c\lambda/K-1/\alpha(n+1,c)), \quad (4.43)$$

其中 $n\geqslant c$, $K>c\geqslant 1$. 据(4.43)和(4.31),我们有

$$\alpha(n+1,K-1)=\frac{2(n+\lambda)+1-\lambda/K-1/\alpha(n,K-1)}{(n+\lambda+1)(n+\lambda+1-\lambda/K)}. \quad (4.44)$$

若引入下列记号

$$p(n)=(2(n+\lambda)+1-\lambda/K)/(n+\lambda+1)\cdot(n+\lambda+1-\lambda/K),$$
$$q(n)=((n+\lambda+1)(n+\lambda+1-\lambda/K))^{-1},$$

则对 $n\geqslant K-1$, $\alpha(n+1,K-1)$ 可以表达为一个连分式

$$\alpha(n+1,K-1)=p(n)-\cfrac{q(n)}{p(n-1)-\cfrac{q(n-1)}{p(n-2)-\ddots-\cfrac{q(K-1)}{\alpha(K-1,K-1)}}}. \quad (4.45)$$

(4.45)把 $\alpha(n+1,K-1)$ 看作是 λ 和 $\alpha(K-1,K-1)$ 的函数. 若对 $n\geqslant K-1$, $\alpha(n,K-1)$ 的值已知, 则对 $n\geqslant K-2$, $\alpha(n,K-2)$ 的值由下列方程作为 $\alpha(K-2,$

$K-2$）的函数来确定：

$$\alpha(n+1,K-2)$$
$$=\frac{1+\alpha(n+1,K-1)(n+(K-2)\lambda/K-1/\alpha(n,K-2))}{n+1+(K-2)\lambda/K}.$$

$$(4.46)$$

一般地，对所有 $c=1,\cdots,K-1$，$n>c$，$\alpha(n,c)$ 的值依赖 n,c,K,λ，且至多依赖值 $\alpha(c,c),\alpha(c+1,c+1),\cdots,\alpha(K-1,K-1)$。据(4.35)和(4.39)可知，当固定 λ 和 $\alpha(c,c)$（其中 $c=1,\cdots,K-1$）的值时，对任意的 n'，n 和 c 我们能确定特征函数 $f(n',n,c)$。因此就能证明下列重要定理：

由公理 4.1—4.5 确定的系统中存在 K 个自由参数。我们可以用 λ 和 $f(0,c,c)$ 来表示，其中

$$c=1,\cdots,K-1. \qquad (4.47)$$

根据(4.47)，我们把公理 4.1—4.5 构成的系统称为归纳逻辑（Niiniluoto 把归纳逻辑看作是研究认知概率的结构性质的逻辑，认为归纳逻辑的基本任务是，确定各个不同的因素如何影响认知主体根据某个证据对假说真值的合理相信度。归纳逻辑不是实际工作的科学家可以运用的一种"方法"，而是提供给哲学家或方法论家的一种工具。例如它为科学哲学提供一种形式工具.）的 K 维系统。后面我们将用下列记号

$$\gamma_c=f(0,c,c)，\text{其中 } c=1,\cdots,K-1.$$

因此 K 维系统除了 λ，还有 $K-1$ 个 gamma 参数。

注意，据(4.42)，我们还可以不用 $\gamma_c(c=1,\cdots,K-1)$，而选择 $f(1,c,c)$（其中 $c=1,\cdots,K-1$）的值作为参数。不用 λ，我们可以据(4.31)选择确定 λ 的 $f(1,K+1,K)$ 作为参数。

利用函数 μ，(4.43) 可以变为

$$\mu(n+1,c+1)[(n+c\lambda/K)/(n+K-c+\lambda)-1/\mu(n,c)]$$
$$=\mu(n+1,c)[(n+1+c\lambda/K)/(n+1+K-c+\lambda)-1/\mu(n+1,c)].\tag{4.48}$$

特别地

$$\mu(n+1,K-1)=\frac{n+\lambda+2}{n+1+\lambda-\lambda/K}\cdot\left(\frac{2(n+\lambda)+1-\lambda/K}{n+1+\lambda}-\frac{1}{\mu(n,K-1)}\right).\tag{4.49}$$

现假设论域U无穷，以便我们可以研究这个特征函数的渐近状态. 当$n\to\infty$时，(4.49)可以用下列方程来近似

$$\mu(n+1,K-1)\mu(n,K-1)-2\mu(n,K-1)+1=0,$$

这个方程的解形如

$$\mu(n,K-1)=(a(n+1)+b)/(an+b),$$

其中a和b是某个常数.因此当$n\to\infty$时，$\mu(n,K-1)\to 1$. 据(4.48)且利用数学归纳法,易证,对所有$c=1,\cdots,K-1$, $\mu(n,c)$渐近地趋于1. 据(4.34),(4.37),(4.38)和(4.39),当$n\to\infty$时，我们有

$$\begin{aligned}&\mu(n,c)\to 1,\\&\alpha(n,c)\to 0,\\&g(n,c)\to 1,\\&h(n,c)\to 0,\\&f(0,n,c)\to 0.\end{aligned}\tag{4.50}$$

据(4.35)和(4.50),当$n\to\infty$时，特征函数$f(n',n,c)$的值趋于相对频率n'/n的极限(若这个极限存在的话),即

$$f(n',n,c)-n'/n$$
$$=[(\mu(n,c)(n'+\lambda/K))/(n+K-c+\lambda)]$$
$$-n'/n\to 0\ (当\ n\to\infty).$$

这个原则被称为 Reichenbach 公理。因此,我们有

公理 4.1—4.5 蕴涵 Reichenbach 公理。 (4.51)

从公理 4.1—4.3 和 Reichenbach 公理,我们有正例证相干原则(除了退化情况):

$$P(Q_i(a_{n+2}),e_n\wedge Q_i(a_{n+1}))>P(Q_i(a_{n+1}),e_n).$$

因此从 (4.51) 推出

正例证相干原则在 K 维系统中成立。 (4.52)

用特征函数,(4.52) 可以如下表示

$$f(n'+1,n+1,c)>f(n',n,c),$$
$$当\ n-c+1\geqslant n'\geqslant 1, \tag{4.53}$$
$$f(1,n+1,c+1)>f(0,n,c). \tag{4.54}$$

据 (4.54),我们有

$$\gamma_c=f(0,c,c)<f(1,c+1,c+1),$$
$$对\ 1\leqslant c<K, \tag{4.55}$$
$$f(0,n,K-1)<f(1,n+1,K)$$
$$=(1+\lambda/K)/(n+1+\lambda). \tag{4.56}$$

特别地,

$$\gamma_{K-1}=f(0,K-1,K-1)<f(1,K,K)=1/K. \tag{4.57}$$

前面我们得到的结果对 K 维系统的参数值作了某些规定。首先据 (4.31),我们有

$$\lambda>-K, \tag{4.58}$$

$$0<f(1,K+1,K)=\frac{K+\lambda}{K(K+\lambda+1)}<1/K. \tag{4.59}$$

$0=f(1,K+1,K)$ 的情况被公理 4.2 排除。(4.42) 和公理 4.1—4.2 蕴涵着 $\gamma_c<1/(K-c)$ 和 $f(1,c,c)<1/c$。据

(4.55)，$\gamma_c < f(1, c+1, c+1) < 1/(c+1)$，而这又蕴涵 $(1-(K-c)/(c+1))/c < f(1,c,c)$。把联立上述结果，有

若 $(K-1)/2 \leqslant c < K$，则

$$0 < \gamma_c < 1/(c+1),$$

$$(1+2c-K)/c(c+1) < f(1,c,c) < 1/c.$$

若 $1 \leqslant c < (K-1)/2$，则

$$0 < \gamma_c < 1/(K-c),$$

$$0 < f(1,c,c) < 1/c. \tag{4.60}$$

我们将在后面看到，(4.60) 没有给出参数值 γ_c 的最准确的上界。考虑归纳概括句的情况说明，选择的 γ 参数不能比 Carnap 选择的值更大。更确切地说，假设 $f(1, K+1, K)$ 的值能在 Carnap 的 λ 系统中用参数 λ_0 来确定，则 (4.31) 推出参数 $\lambda = \lambda_0$。对应于 $\lambda > 0$ 的每一选择，据 (4.29)，Carnap 的 γ_c 为

$$\gamma_c = \frac{\lambda/K}{c+\lambda}, \quad c = 1, \cdots, K-1. \tag{4.61}$$

（注意，若 $-K < \lambda \leqslant 0$，则 γ_c 不能选取这样的值。）

从前面的一些结果可见，对 $n' > 0$，$f(n', n, c)$ 以及 $f(0,n,c)$ 和 $g(n,c)$ 可以用参数 λ 和 γ_i（其中 $i = c, \cdots, K-1$）来确定，因此它们不依赖 $\gamma_1, \cdots, \gamma_{c-1}$。下面对 Niiniluoto 的 K 个参数的作用给出更详细的描述。当 $1 \leqslant c < K$，据 (4.42)，(4.43) 和 (4.39)，有

$$\begin{aligned} & f(0,c,c)/f(0,c+1,c) \\ = & \gamma_c(K-c)/(1-(c+c\lambda/K+1)\alpha(c+1,c)) \\ = & \gamma_c(K-c)/[\alpha(c+1,c+1)(1/\alpha(c,c)-c-c\lambda/K)] \\ = & ((c+1)(1-(K-c)\gamma_c))/(c(1-(K-c-1)\gamma_{c+1})). \end{aligned}$$

因此，

$$f(0,c+1,c) = \frac{1-(K-c-1)\gamma_{c+1}}{1-(K-c)\gamma_c}$$

$$\cdot \frac{c}{c+1}\gamma_c. \tag{4.62}$$

对 $c=K-1$，我们有

$$f(0,K,K-1)=\frac{K-1}{K}\cdot\frac{\gamma_{K-1}}{1-\gamma_{K-1}}. \tag{4.63}$$

同理可证,对 $1\leqslant c<K$ 且 $j\geqslant 0$，我们得

$$f(0,c+j+1,c)$$

$$=\frac{(c+j+c\lambda/K)(1-(K-c-1)f(0,c+j+1,c+1))}{(c+j+1+(c+1)\lambda/K)(1-(K-c)f(0,c+j,c))}$$

$$\cdot f(0,c+j,c). \tag{4.64}$$

据(4.62)和(4.63),且利用归纳法，我们有

令 $1\leqslant c<K$，则对 $n'\geqslant 0$，$f(n',c+1,c)$ 是 γ_c 和 γ_{c+1} 的函数，但独立于其他参数. 对 $1<j\leqslant K-c-1$ 且 $n'\geqslant 0$，$f(n',c+j,c)$ 是 λ 和 $\gamma_c,\cdots,\gamma_{c+j}$ 的函数。(4.65)

(4.65)精确地说明,在什么条件下,每一参数对单称归纳推理是相关的.

从(4.62)和(4.64)我们可以看到 γ 参数的影响. 据(4.62)，$f(0,c+1,c)$ 随 γ_c 的递增而递增,但随 γ_{c+1} 的递增而递减. 据(4.64)和数学归纳法，能更一般地证明，对 $n\geqslant c$，$f(0,n,c)$ 随 γ_c 递增而递增，但随 $f(0,n,c)$ 所依赖的所有其他参数的递增而递减.

对 $n\geqslant c$，$f(0,n,c)$ 随 γ_c 的递增而递增但随其他有关参数的递增而递减,因此对 $n'>0$，$f(n',n,c)$ 和 $g(n,c)$ 随 γ_c 的递增而递减,但随其他有关参数的递增而递增. (4.66)

(4.66)说明：选择的 γ_c 越小,概率 $g(n,c)$ 越大，即期望下一个体属于已被证据例证的 c 个 Q 类之一的可能性也越大.

从§1已知，语言 $\mathscr{L}$ 的构元是在 $\mathscr{L}$ 的概括句中最强的概括句,对每一 Q 类 Q_i（其中 $i=1,\cdots,K$)，这些构元规定

Q_i 是否在论域 U 中被例证. 这也说明了 w 个 $Q_{i_1}, \cdots, Q_{i_w}$ 在 U 中被例证的构元 C_w 有下列形式

$$\bigwedge_{j=1}^{w} (\exists x) Q_{i_j}(x) \wedge (x) \left(\bigvee_{j=1}^{w} Q_{i_j}(x) \right), \qquad (4.67)$$

这里的 w 称为 C_w 的宽度. 据 [11], $\mathscr{L}$ 中每一概括句都可以用有穷个构元的析取范式来表达. 因为 $\mathscr{L}$ 的构元互不相容, 所以 $\mathscr{L}$ 中任意概括句的先验(或后验)概率都等于这个析取范式中的构元的先验(或后验)概率之和. 因此归纳概括推理问题就归结为确定构元的先验和后验概率的问题.

下列原则可以看作是归纳概括句理论的恰当性的基本条件.

若 e_n^c 与 C_w 相容且 c 固定, 则

$$P(C_w, e_n^c) \to 1, \text{ 若 } w = c \text{ 且 } n \to \infty,$$
$$P(C_w, e_n^c) \to 0, \text{ 若 } w > c \text{ 且 } n \to \infty, \qquad (4.68)$$

这里的"e_n^c"描述了具有 n 个个体且形成 c 个个体类的样本.

下面我们在 K 维系统中确定构元的概率. 首先注意到

$$P(e_n^c) = f(0,0,0) f(0,1,1) \cdots f(0, c-1, c-1)$$
$$= \frac{1}{K} \cdot \gamma_1 \cdots \gamma_{c-1}. \qquad (4.69)$$

更一般地, 若在 n_i (其中 $i = 1, \cdots, K$) 中非零数为 $n_1, \cdots, n_c$, 且令 $N_0 = 0$, 对 $j = 1, \cdots, c$, 令 $N_j = n_1 + \cdots + n_j$, 则

$$P(e_n^c) = f(0,0,0) f(1,1,1) \cdots f(n_1 - 1, n_1 - 1, 1)$$
$$\cdot f(0, n_1, 1) f(1, n_1 + 1, 2) \cdots f(n_2 - 1, n_1 + n_2 - 1, 2)$$
$$\cdots f(0, n_1 + \cdots + n_{c-1}, c-1)$$
$$\cdot f(1, n_1 + \cdots + n_{c-1} + 1, c) \cdots f(n_c - 1, n - 1, c)$$
$$= \prod_{i=0}^{c-1} f(0, N_i, i) \prod_{j=0}^{c-1} \prod_{k=1}^{n_{j+1}-1} f(k, N_j + k, j+1), \qquad (4.70)$$

似然度 $P(e_n^c, C_w)$ 由下列概率来确定

$$P(Q_i(a_{n+1}), e_n^c \wedge C_w). \tag{4.71}$$

在 Hintikka 的 α-λ 系统中,(4.71) 由下列公式来定义

$$(n_i + \lambda(w)/w)/(n + \lambda(w)), \tag{4.72}$$

其中 $\lambda(w)$ 的值可以依赖 w 但不依赖 c。换句话说，相对化特征函数 (4.71) 的值与 w 个 Q 类的论域的 Carnap 特征函数的值相对应。

通过对上述有关论证相对化，可以看到，对 $w = 1, \cdots, K$，存在函数 $f_w(n', n, c)$ (其中 $c \leqslant w$)，使得 (4.71) 表达为 n', n 和 c 的函数，但独立于数 n_j (其中 $j \neq i$)，而且函数 f_w 完全由数 $f_w(1, w+1, w)$ 和下列 $w-1$ 个数确定，

$$\delta_c^w = f_w(0, c, c), \text{ 对 } c = 1, \cdots, w-1. \tag{4.73}$$

因此，从 Bayes 定理和 (4.69)，我们有

$$\begin{aligned} P(C_K) &= P(e_K^K)P(C_K, e_K^K)/P(e_K^K, C_K) \\ &= P(e_K^K)/P(e_K^K, C_K) \\ &= \left(\frac{1}{K} \cdot \gamma_1 \cdots \gamma_{K-1}\right) \Big/ \left(\frac{1}{K} \cdot \delta_1^K \cdots \delta_{K-1}^K\right). \end{aligned} \tag{4.74}$$

若对所有 $i \neq c$，令 $\gamma_i = \delta_i^K$，则从公理 4.1 推出

$$\gamma_c \leqslant \delta_c^K, \text{ 对 } c = 1, \cdots, K-1, \tag{4.75}$$

因此

$$P(C_K) = 1 \Longleftrightarrow \gamma_c = \delta_c^K, \text{ 对所有 } c = 1, \cdots, K-1. \tag{4.76}$$

因为 Carnap 的 λ 系统是 K 维系统的一个特例，所以当 γ_c 等于它的 Carnap 值时，$P(C_K) = 1$。因此 δ_c^K 必须等于这些 Carnap 值。所以据 (4.61)，

$$\delta_c^K = \frac{\lambda/K}{c + \lambda}, \text{ 对 } c = 1, \cdots, K-1. \tag{4.77}$$

因此，

$$P(C_K)=(\gamma_1\cdots\gamma_{K-1})/(\delta_1^K\cdots\delta_{K-1}^K)$$
$$=(1+\lambda)(2+\lambda)\cdots(K-1+\lambda)\gamma_1\cdots\gamma_{K-1}/(\lambda/K)^{K-1}. \quad (4.78)$$

这里的(4.75)和(4.77)与前面的(4.60)相比，给出了关于γ参数的更精确的上界，即给定 $\lambda>0$，对 $c=1,\cdots,K-1$，参数 γ_c 的可容许范围是

$$0<\gamma_c\leqslant(\lambda/K)/(c+\lambda). \quad (4.79)$$

(4.75)—(4.79) 刻划了 Niiniluoto 的K维系统和Carnap的λ系统之间的关系。Carnap 的λ系统是K维系统当 $P(C_K)=1$ 的特例。换句话说，$P(C_K)=1$ 在K维系统成立当且仅当所有的参数 γ_c（其中 $c=1,\cdots,K-1$）有它们的 Carnap 值。在所有其他的情况下，至少有一个全称事实概括句有非零的先验概率。

为了概括上述论证，我们假设论域U的$K-w$个Q类是空的，来令 P_w 是相对化P得到的概率测度。假设 C_w 和 e_w^w 与关于U的这个假设相容，则

$$P_w(C_w)=P_w(e_w^w)P_w(C_w,e_w^w)/P_w(e_w^w,C_w)$$
$$=P_w(e_w^w)/P_w(e_w^w,C_w).$$

如上所述，存在唯一的选择 $P_w(Q_i(a_{c+1}),e_c^c)$ 的方法，使得 $P_w(C_w)=1$，$P_w(Q_i(a_{c+1}),e_c^c)=\delta_c^w$。又因为 Carnap 选择 $P_w(Q_i(a_{c+1}),e_c^c)$ 来保证 $P_w(C_w)=1$，所以存在数 $\lambda(w)$ $(w=1,\cdots,K)$ 使得

$$\delta_c^w=f_w(0,c,c)=(\lambda(w)/w)/(c+\lambda(w)), \quad (4.80)$$

而且 $\lambda(w)$ 必须满足方程

$$\lambda(w)=\frac{w\cdot f_w(1,w+1,w)}{1-w\cdot f_w(1,w+1,w)}-w, \quad (4.81)$$

其中 $f_w(1,w+1,w)=P(Q_i(a_{w+2}),e_{w+1}^w\wedge C_w)$，只要 $n_i=$

1. 当 $\lambda(K)=\lambda$ 时，(4.40) 给出

$$f_w(1,w+1,w)=(1+\lambda/K)/(w+1+w\lambda/K). \quad (4.82)$$

把 (4.82) 代入 (4.81)，我们有

$$\lambda(w)=w\lambda/K, \quad (4.83)$$

即相对每一 λ，$\lambda(w)/w$ 是独立于 w 的常数项。据 (4.80)，我们有

$$\delta_c^w=(\lambda/K)/(c+w\lambda/K). \quad (4.84)$$

因此 δ_c^w 和 δ_c^K 的关系是

$$\delta_c^w=\frac{c+\lambda}{c+w\lambda/K}\cdot\delta_c^K. \quad (4.85)$$

这些结果说明，在 K 维系统中构元的似然度 $P(e_n^c,C_w)$ 确定为 $n,n_1,\cdots,n_c,w$ 和 λ 的函数，但这些似然度独立于 γ 参数。在这个系统的 K 个参数中，只有 λ 与这些似然度相关. 为了得到关于似然度的公式，首先据相对 C_w 的 (4.31)，有

$$f_w(n',n,w)=\frac{n'+\lambda(w)/w}{n+\lambda(w)}=\frac{n'+\lambda/K}{n+\lambda w/K}.$$

据 (4.84) 和数学归纳法，再用 (4.64) 的相对化形式，我们有，对所有 $n\geqslant c$

$$f_w(0,n,c)=(\lambda/K)/(n+w\lambda/K). \quad (4.86)$$

据 (4.41) 的相对化形式，我们有

$$f_w(n',n,c)=\frac{n'+\lambda/K}{n+c\lambda/K}\left(1-\frac{(w-c)\lambda/K}{n+w\lambda/K}\right).$$

而这蕴涵：对所有 $n'\geqslant 0$，

$$f_w(n',n,c)=(n'+\lambda/K)/(n+w\lambda/K). \quad (4.87)$$

因此似然度就可以从下列公式中得到：

$$P(e_n^c,C_w)=\prod_{i=1}^{c}\prod_{j=0}^{n_i-1}(j+\lambda/K)\Big/\prod_{k=1}^{n-1}(k+w\lambda/K)$$

$$=\left[\Gamma(w\lambda/K)\prod_{i=1}^{c}\Gamma(n_i+\lambda/K)\right]\Big/$$

$$[\Gamma(\lambda/K)^c \cdot \Gamma(n + w\lambda/K)], \qquad (4.88)$$

其中Γ是 gamma 函数。

(4.88)说明，在K维系统中确定似然度就像在 Hintikka 的α-λ 系统 ($\lambda(w)/w$ 是一常数)那样。α-λ 系统中的这些元素只能是K维系统中的元素。例如，令 $\lambda(w)/w = 1$ 而得到的 Hintikka 的"概括的联合系统"属于K维系统。更一般地，我们有如下推论：

K维系统和 α-λ 系统的交正好包含 α-λ 系统中满足 $\lambda(w) = aw$ (其中 $a > 0$ 是一常数) 的那些元素。 (4.89)

因此我们可以说 Niiniluoto 的K维系统本质上是 Hintikka 的"联合系统"的概括，而后者是 Carnap 的确证函数 c^* 的概括。

下面我们来研究关于构元的先验概率和后验概率的某些结果。从全概定理知

$$P(e_n^w) = \sum_{i=0}^{K-w} \binom{K-w}{i} P(C_{w+i}) P(e_n^w, C_{w+i}),$$

特别是

$$\begin{aligned} P(e_{K-1}^{K-1}) &= P(C_{K-1})P(e_{K-1}^{K-1}, C_{K-1}) \\ &\quad + P(C_K)P(e_{K-1}^{K-1}, C_K), \\ P(C_{K-1}, e_{K-1}^{K-1}) &= P(C_{K-1})P(e_{K-1}^{K-1}, C_{K-1})/P(e_{K-1}^{K-1}) \\ &= (P(e_{K-1}^{K-1}) - P(C_K)P(e_{K-1}^{K-1}, C_K))/P(e_{K-1}^{K-1}) \\ &= 1 - \gamma_1 \cdots \gamma_{K-1} \cdot \frac{1}{K-1} \\ &\quad \cdot \delta_1^K \cdots \delta_{K-2}^K / \delta_1^K \cdots \delta_{K-1}^K \cdot \frac{1}{K-1} \\ &\quad \cdot \gamma_1 \cdots \gamma_{K-2}. \end{aligned} \qquad (4.90)$$

因此

$$P(C_{K-1}, e_{K-1}^{K-1}) = 1 - \gamma_{K-1}/\delta_{K-1}^{K}$$

$$= 1 - \frac{(K-1+\lambda)}{\lambda/K} \cdot \gamma_{K-1}. \tag{4.91}$$

据 Bayes 定理,我们就能得到 C_{K-1} 的先验概率

$$P(C_{K-1}) = \frac{(K-1)}{K} \cdot \frac{\gamma_1 \cdots \gamma_{K-2}}{\delta_1^{K-1} \cdots \delta_{K-2}^{K-1}} \left(1 - \frac{\gamma_{K-1}}{\delta_{K-1}^{K}}\right). \tag{4.92}$$

对 $w = K-1, K-2, \cdots, 2$, 我们可以用类似的方法, 从 (4.90) 相继得到 $P(C_w, e_w^w)$ 和 $P(C_w)$.

下面的结果说明了在归纳概括推理中 γ 参数的作用. 对 $w < K$,

$$对\ i = w, \cdots, K-1, 若\ \gamma_i = \delta_i^K, P(C_w) = 0, \tag{4.93}$$

$$当\ \gamma_w\ 递增时, P(C_w)\ 递减, \tag{4.94}$$

$$P(C_w, e_w^w) \to 1\ (当\ \gamma_w \to 0). \tag{4.95}$$

因此参数 γ_w 可以看作是断定宽度为 w 的构元 C_w 为真的谨慎度的指标. 在这个方面, 这些 γ 参数的特征与 Hintikka 的 α-λ 系统中的参数 α 的特征类似. 但 Hintikka 的 α 是关于 $C_w, w < K$, 为真的全部谨慎度的指标,而在 K 维系统中,对每一宽度 $w < K$, 存在一个分开的谨慎度指标. 因此对于后者,选择构元的先验概率就自由得多(注意: 在 α-λ 系统和 K 维系统中,参数 λ 以相同的方式刻划单称归纳推理).

原则上,对 $c \leqslant w$ 和 $n \geqslant c$, 形如 $P(C_w, e_n^c)$ 的后验概率能从 (4.90) 得到. 例如,下列结果是对 (4.91) 的概括: 对 $c = 1, \cdots, K-1$,

$$P(C_{K-1}, e_c^c) = 1 - \gamma_c \cdots \gamma_{K-1}/\delta_c^K \cdots \delta_{K-1}^K, \tag{4.96}$$

在无穷论域 U 中, $P(C_w, e_n^w)$ 可以用 g 表示:

$$P(C_w, e_n^w) = \prod_{i=0}^{\infty} g(n+i, w). \tag{4.97}$$

当 $w = K-1$ 且 λ 是 K 的倍数时,从 (4.97)

$$P(C_{K-1}, e_n^{K-1}) = 1 - \frac{n+\lambda}{\lambda/K} \cdot f(0, n, K-1). \quad (4.98)$$

(当 $n = K-1$ 时参考(4.91))。 因此在基数 $m+n > \lambda/K$ 的有穷论域 U 中,我们有

$$P(C_{K-1}, e_n^{K-1}) = \prod_{i=0}^{m} g(n+i, K-1)$$

$$= g(n, K-1) - (1 - g(n, K-1))$$

$$\cdot \frac{K}{\lambda}(n+(K-1)\lambda/K)\left(m - \frac{(n+\lambda)!}{(n+\lambda-\lambda/K)!}\right.$$

$$\left.\cdot \sum_{i=1}^{m} \frac{n+\lambda-\lambda/K+m+1}{n+\lambda+m+1}\right). \quad (4.99)$$

可以证明,当 $n\to\infty$ 时,(4.98)趋于1. 更一般地,据数学归纳法,

若选择参数 γ_c (其中 $c = w, \cdots, K-1$) 比它的 Carnap 值 δ_c^K 更小,则当 w 固定且 $n\to\infty$ 时,

$$P(C_w, e_n^w) \to 1. \quad (4.100)$$

这说明当选择参数小于 Carnap 值,基本恰当条件(4.68)在 K 维系统中成立. 对(4.68)唯一的例外是 Carnap 的条件,在其他所有的条件下,全称事实概括句有非零的先验概率,并且具有逐渐趋于1的后验概率。

第五章　条件句概率逻辑

本世纪 60 年代，E. W. Adams[12] 提出了条件句概率逻辑。主要思想是用概率来研究条件句逻辑，即用推理的“合理性”（reasonableness）标准来代替用真值条件陈述的推理的“有效性”标准，用“高概率”来代替替经典逻辑的“有效性”定义中“真”的概念。所谓“合理性”标准就是要求一合理推理的前提的高概率必须确保结论也有高概率。Adams 的条件句逻辑在某些方面是 Carnap 归纳逻辑的精致化。我们认为他的贡献在于在发展 Carnap 思想的同时，建立了概率逻辑的形式语法系统，并且证明了概率后承集和合理后承集重合的完全性定理，第一次使概率逻辑具有完备的形式。

§1　基本概念

定义 5.1　（1）一公式是任一表达式使得满足下列条件之一：

（i）原子公式：“T”，“⊥”（分别表示逻辑真和逻辑假）和所有带下标或不带下标的原子句：$p, q, r, p_1, q_1, r_1, \cdots$。

（ii）真值函数公式：所有的原子句和用通常意义的真值联结词“∧”，“∨”，“¬”构造的表达式。

（iii）条件句公式：所有形如 $\varphi \to \psi$ 的公式，其中 φ 和 ψ 是真值函数公式。

（2）令 α 是包括“T”和“⊥”的原子公式集，α 的语言

$$\mathscr{L}(\alpha)=\alpha\cup\mathscr{L}^*(\alpha)\cup\mathscr{L}'(\alpha),$$

其中 $\mathscr{L}^*(\alpha)$ 是由 α 的原子公式用“∧”、“∨”、“┐”布尔组合构造起来的真值函数公式集，$\mathscr{L}'(\alpha)$ 是由 $\mathscr{L}^*(\alpha)\cup\alpha$ 的元素通过“→”作为主联结词构造起来的条件句公式集. 语言 $\mathscr{L}$ 是一公式集，由某个 α 以上述方法生成的. 称一语言 $\mathscr{L}$ 有穷，若生成 $\mathscr{L}$ 的原子公式集有穷.

为了方便，我们说一公式或一公式集 S 重言蕴涵公式 A（即对所有真值赋值 f，若对所有 $B\in S$，$f(B)$ 为真，则 $f(A)$ 为真，记作 $S\vDash A$），甚至 S 和 A 可以包括条件句公式. 在运用重言式和重言蕴涵的概念于条件句公式时，我们把它看作是对应的实质蕴涵句. 例如说 $\varphi\to\psi\vDash\psi'$ 就是说 $\neg\varphi\vee\psi$ 重言蕴涵 ψ'.

定义 5.2 令 $A\in\mathscr{L}$，$\varphi,\psi\in\mathscr{L}^*$

(1) 若 $A=\psi$，则 $\sim A=\neg\psi$；若 $A=\varphi\to\psi$，则 $\sim A=\varphi\to\neg\psi$.

(2) 若 $A=\psi$，则 $\mathrm{Ant}(A)=\top$（$\mathrm{Ant}(A)$ 表示 A 的前件），$\mathrm{Cons}(A)=\psi$ 且 $\mathrm{Cond}(A)=\top\to\psi$. 若 $A=\varphi\to\psi$，则 $\mathrm{Ant}(A)=\varphi$，$\mathrm{Cons}(A)=\psi$ 且 $\mathrm{Cond}(A)=\varphi\to\psi$.

下列概念只限于有穷语言. 我们称 $\mathrm{SD}\mathscr{L}$ 是 $\mathscr{L}$ 的状态描述集，若 $\mathscr{L}$ 有穷且 $\mathrm{SD}\mathscr{L}$ 具有下列性质：(i) $\mathrm{SD}\mathscr{L}\subseteq_f\mathscr{L}^*$（$A\subseteq_f B$ 表示 A 是 B 的有穷子集，下同），且 $\mathrm{SD}\mathscr{L}$ 重言一致（即 $\mathrm{SD}\mathscr{L}$ 的所有元的析取重言等价 $\top$）. (ii) $\mathrm{SD}\mathscr{L}$ 的任意两不同元素 α，β（表示两不同的状态描述）是重言不一致的（即 $\alpha\wedge\beta$ 重言等价 $\bot$，这里的 $\top$ 和 $\bot$ 分别表示重言式和重言式的否定），$\mathscr{L}^*$ 的不重言等价 $\bot$ 的每一真值函数公式重言等价由 $\mathrm{SD}\mathscr{L}$ 的若干元素构成的（唯一的）析取. 有穷语言的 $\mathrm{SD}\mathscr{L}$ 以类似 Carnap 的方法构造，若 $\mathscr{L}=\{p_1,\cdots,$

$p_n\}$，则 $\mathrm{SD}\mathscr{L}=\left\{\bigwedge_{i=1}^{n} p_i' : p_i' \text{ 是 } p_i \text{ 或 } \neg p_i\right\}$. 对有穷语言的状态描述集的赋值由下列定义规定：

定义 5.3 令 $\alpha\in \mathrm{SD}\mathscr{L}$，$S\subseteq_f \mathscr{L}$，$A\in\mathscr{L}$，$\mathrm{Ant}(A)=\varphi$，$\mathrm{Cons}(A)=\psi$，

(1) $\mathrm{SD}(A)=\{\alpha\in \mathrm{SD}\mathscr{L} : \alpha\models\varphi\}$，$\boldsymbol{SD}(A)=\{\alpha\in \mathrm{SD}\mathscr{L} : \alpha\models\varphi\wedge\neg\psi\}$.

(2) $\mathrm{SD}(S)=\bigcup_{B\in S}\mathrm{SD}(B)$，$\boldsymbol{SD}(S)=\bigcup_{B\in S}\boldsymbol{SD}(B)$.

(3) S 是零的 (null) $\Leftrightarrow \mathrm{SD}(S)=\boldsymbol{SD}(S)$.

注意，$\mathrm{SD}(A)$ 类似于 Carnap 的 $\mathrm{RA}(A)$. $\boldsymbol{SD}(A)=\{\alpha\in \mathrm{SD}\mathscr{L} : \alpha\not\models A\}$，因为 $\alpha\in\boldsymbol{SD}(A)\Leftrightarrow\alpha\not\models\neg(\varphi\wedge\neg\psi)\Leftrightarrow \alpha\not\models A$. 此外，总有 $\boldsymbol{SD}(A)\subseteq\mathrm{SD}(A)$，$\mathrm{SD}(A)=\boldsymbol{SD}(A)\cup\boldsymbol{SD}(\sim A)$，$\{\alpha\in\mathrm{SD}\mathscr{L} : \alpha\models\varphi\wedge\psi\}=\mathrm{SD}(A)-\boldsymbol{SD}(A)$.

定义 5.4 令 $\mathscr{L}$ 是一语言，$\mathscr{L}$ 的概率函数 P 是以 $\mathscr{L}$ 为定义域的实值函数使得对所有 $\varphi,\psi\in\mathscr{L}^*$，则

(1) $0\leqslant P(\varphi)\leqslant 1$ 且 $P(\top)=1$.

(2) 若 $\varphi\models\psi$，则 $P(\varphi)\leqslant P(\psi)$.

(3) 若 $\varphi\models\neg\psi$，则 $P(\varphi\vee\psi)=P(\varphi)+P(\psi)$.

(4) $$P(\varphi\to\psi)=\begin{cases}P(\varphi\wedge\psi)/P(\varphi), & \text{若 } P(\varphi)\neq 0,\\ 1, & \text{否则}.\end{cases}$$

注意，至少对 (2)，φ,ψ 必须是真值函数公式，而不是条件句. 例如，令 $\varphi=p$，$\psi=\neg p\to\bot$，则 $p\models\neg p\to\bot$，但 $P(p)\not\leqslant P(\neg p\to\bot)$，只要取一概率函数 P 使得 $P(p)=\frac{1}{2}$，则 $P(p)\leqslant P(\neg p\to\bot)$ 不成立.

定理 5.1

(1) $P(\neg\varphi)=1-P(\varphi)$.

(2) $P(\varphi) = P(\varphi \wedge \psi) + P(\varphi \wedge \neg\psi)$.

(3) $P(\varphi \to \neg\psi) = 1 - P(\varphi \to \psi)$，若 $P(\varphi) \neq 0$.

(4) $P(\varphi \vee \psi) = P(\varphi) + P(\psi) - P(\varphi \wedge \psi)$.

(5) $P(\varphi \vee \psi) = 0 \Longleftrightarrow P(\varphi) = 0$ 且 $P(\psi) = 0$.

(6) $P(\varphi \vee \psi) = P(\varphi) + P(\psi \wedge \neg\varphi)$.

(7) $P(\varphi \vee \gamma \to \gamma) \geqslant P(\varphi \vee \psi \to \psi) \cdot P(\psi \vee \gamma \to \gamma)$.

(8) 若 $P(\varphi \vee \gamma \to \gamma) > 0$ 且 $P(\psi \vee \gamma \to \gamma) > 0$，则 $P(\varphi \vee \psi \vee \gamma \to \gamma) > 0$.

(9) 若 $P(\varphi \vee \psi \to \psi) > 0$ 或 $P(\varphi \vee \gamma \to \gamma) > 0$，则 $P(\varphi \vee \psi \vee \gamma \to \psi \vee \gamma) > 0$.

证明留给读者做练习。

定义 5.5 $\mathscr{L}$ 的齐一概率函数序列是 $\mathscr{L}$ 的无穷的概率函数序列 $P_1, P_2, \cdots$ 使得对所有 $A \in \mathscr{L}$，$\lim\limits_{n \to \infty} P_n(A)$ 存在。

定义 5.6 令 $\mathscr{L}$ 是语言，$S \subseteq \mathscr{L}$，$A \in \mathscr{L}$，我们有

(1) A 是 S 的合理后承(记为 $S \Vdash A$) $\Longleftrightarrow$ 对所有 $\varepsilon > 0$，存在 $\delta > 0$，使得对 $\mathscr{L}$ 的所有概率函数 P，若对所有 $B \in S$，$P(B) > 1 - \delta$，则 $P(A) > 1 - \varepsilon$。

(2) A 是 S 的严格后承(记为 $S \vdash A$) $\Longleftrightarrow$ 对 $\mathscr{L}$ 的所有概率函数，若对所有 $B \in S$，$P(B) = 1$，则 $P(A) = 1$。

注意，(1) 是 Adams 系统的主要语义概念，(2) 是它的极限情况。

定理 5.2 令 $\mathscr{L}$ 是一语言，$S \subseteq \mathscr{L}$，$A \in \mathscr{L}$，则有

(1) $S \vdash A \Longleftrightarrow S \models A$ (即 A 是 S 的重言后承)。

(2) 若 $S \Vdash A$，则 $S \vdash A$。

(3) 若 S，$S' \subseteq_f \mathscr{L}$，则

(i) 若 $A \in S$，则 $S \Vdash A$。

(ii) 若对所有 $B \in S$，$S' \Vdash B$，且 $S \Vdash A$，则 $S' \Vdash A$。

(iii) 如果 S' 和 A' 分别是从 S 和 A 用真值函数公式 φ 处处代换原子公式 α（不是⊤和⊥）得到的，那么，若 $S \Vdash A$，则 $S' \Vdash A'$。

证明：(1) 若 $S \not\models A$，则对 $\mathscr{L}$ 的所有公式，存在一真值赋值 f（其中对条件句公式赋真值时把它们看作是对应的实质蕴涵句）使得对所有 $B \in S, f(B)$ 为真但 $f(A)$ 为假。现定义 $\mathscr{L}$ 上一函数 P 使得对所有 $C \in \mathscr{L}$，

$$P(C) = \begin{cases} 1, & \text{若 } f(C) \text{ 为真}, \\ 0, & \text{若 } f(C) \text{ 为假}. \end{cases}$$

易证 P 满足定义 5.4，因此构成 $\mathscr{L}$ 上一概率函数。根据 P 的构造，对所有 $B \in S, P(B) = 1$ 但 $P(A) = 0$，因此 $S \not\Vdash A$。

若 $S \models A$，则据古典命题演算的完全性定理，存在 S 的有穷子集 $S' = \{B_1, \cdots, B_n\}$ 使得 $S' \models A$。令 $B_1', \cdots, B_n'$，A' 分别是和 $B_1, \cdots, B_n$，A 相对应的实质蕴涵句，因此 $B_1' \wedge \cdots \wedge B_n' \to A$ 是重言式，且对任意概率函数，它的概率为 1。从而对于任意概率函数 P（使得对 $i = 1, \cdots, n$，$P(B_i) = 1$），赋值 $P(A') = 1$。但从概率函数的定义，若 C 是任一公式，C' 是它对应的实质蕴涵句（若 $C \in \mathscr{L}^*$，则 $C' = C$），则 $P(C) = 1 \Longleftrightarrow P(C') = 1$。证明如下：

令 $C = \varphi \to \psi$，$C' = \varphi \supset \psi$（"$\supset$"是实质蕴涵号）。先证"$\Leftarrow$"。设 $P(\varphi) \neq 0$，否则显然。若 $P(C') = 1$，则 $1 = P(\neg\varphi \vee \psi) = 1 - P(\varphi) + P(\psi) - P(\neg\varphi \wedge \psi)$，因此 $P(\psi) = P(\varphi) + P(\neg\varphi \wedge \psi)$，则据定理 5.1，$P(\psi \wedge \varphi) + P(\psi \wedge \neg\varphi) = P(\varphi) + P(\neg\varphi \wedge \psi)$，即 $P(\varphi \wedge \psi) = P(\varphi)$，因此 $P(\varphi \wedge \psi)/P(\varphi) = P(\varphi \to \psi) = 1$。"$\Rightarrow$"，若 $P(\varphi) = 0$，则 $P(\neg\varphi) = 1$，因为 $\neg\varphi \models \varphi \supset \psi$，所以 $P(\varphi \supset \psi) = P(C') = 1$。若 $P(\varphi) \neq 0$，则 $1 = P(C) = P(\varphi \wedge \psi)/P(\varphi)$，所以 $P(\varphi \wedge \psi) = P(\varphi)$，$P(\varphi \supset \psi) = P(\neg\varphi \vee \psi) =$

$1-P(\varphi)+P(\psi)-P(\neg\varphi\wedge\psi)=1-P(\varphi\wedge\psi)+P(\psi)-P(\neg\varphi\wedge\psi)=1$.

因此若对 $i=1,\cdots,n,P(B_i)=1$，则 $P(A)=1$，亦即 $S'\vdash A$，从而 $S\vdash A$，这个定理实际上说明只取 0，1 为值的 P 函数与真值函数等价.

(2) 若 $S\nvdash A$，则据 (1)，存在对 $\mathscr{L}$ 的真值赋值，在该赋值下，S 的所有公式为真但 A 为假. 令 (1) 的证明中的概率函数 P 使得对所有 $B\in S$，$P(B)=1$ 但 $P(A)=0$，因此 $S\Vdash\!\!\!\!/\, A$. 注意 (2) 的逆不成立，例如 $\neg p\rightarrow\bot$ 是 p 的重言后承，但不是合理后承，我们证明如下：

取某个 $\varepsilon>0$，对任意 $\delta>0$，设 $P(p)>1-\delta$，即 $P(\neg p)<\delta$，则据定义 5.4(4)，

$$P(\neg p\rightarrow\bot)=\begin{cases}P(\neg p\wedge\bot)/P(\neg p), & 若\ P(\neg p)\neq 0,\\ 1, & 否则.\end{cases}$$

因为 $\bot$ 是 $\neg p\wedge\bot$ 的重言后承，所以据定义 5.4(2)，$P(\neg p\wedge\bot)\leqslant P(\bot)$，而据定义 5.4(1) 和定理 5.1，$P(\bot)=0$，因此 $P(\neg p\wedge\bot)=0$，所以当 $P(\neg p)\neq 0$，即 $P(p)\neq 1$ 时，$P(\neg p\rightarrow\bot)=0\not>1-\varepsilon$，所以不满足定义 5.6 (1)，因此 $p\Vdash\!\!\!\!/\,\neg p\rightarrow\bot$. (2) 的逆不成立说明合理后承概念的确是重言后承概念的一种真扩张.

(3) (i) 显然，若 $A\in S$，则据定义 5.6 (1)，只要取 $\delta=\varepsilon$，就有 $S\Vdash A$. 注意 (i) 的逆不成立，令 $S=\{\varphi\wedge\varphi\}$，显然有 $S\Vdash\varphi$，但 $\varphi\notin S$.

(ii) 令 $S,S'\subseteq\mathscr{L}$，假设对所有 $B\in S,S'\Vdash B$(S 是 S' 的合理后承集)，且 $S\Vdash A$，则对任意 $\varepsilon>0$，存在 $\delta>0$ 使得若对所有 $B\in S,P(B)>1-\delta$,则 $P(A)>1-\varepsilon$ (这里的

P是$\mathscr{L}$的任意概率函数).因为对所有$B \in S, S' \Vdash B$，所以对所有B，存在δ_B使得若对所有$C' \in S', P(C) > 1 - \delta_B$，则$P(B) > 1 - \delta$。又因为$S$有穷，所以对所有$B \in S$，存在极小$\delta_B > 0$。令$\delta_0$是这个极小值。显然，若对所有$C \in S'$，$P(C) > 1 - \delta_0$，则对所有$B \in S$，$P(B) > 1 - \delta$，因此$P(A) > 1 - \varepsilon$，从而$S' \Vdash A$((ii)说明合理后承的性质是可传递的)。

(iii) 这个定理可直接从下列事实得到：对$\mathscr{L}$的任一概率函数P'，有可能构造$\mathscr{L}$的另一概率函数P使得对所有$B, B' \in \mathscr{L}, P(B) = P'(B')$，其中$B'$是从$B$用$\varphi$处处代换$\alpha$得到的公式. P的这种构造是基本的:任取概率函数P'，直接构造函数P，

$$P(B') = P(B).$$

易证P也是概率函数. 假定P已构造完毕,就可直接推出:若$S' \nVdash A'$，则$S \nVdash A$。因为若存在某个$\varepsilon > 0$，对所有$\delta > 0$，存在一概率函数P'，使得对所有$B' \in S'$，$P'(B') > 1 - \delta$但$P'(A') \leqslant 1 - \varepsilon$，则对所有$B \in S$，$P(B) > 1 - \delta$但$P(A) \leqslant 1 - \varepsilon$，因此$S \nVdash A$((iii)说明合理后承关系对代换保持)。

定理5.2 (3) 说明合理后承关系（至少用于有穷前提集时)至少具有演绎关系的某些极小性质.例如合理后承的合理后承自身也是合理后承,但合理后承关系不是演绎关系,若这个关系的前域包括无穷公式集。这可以从它不满足下列紧致性条件看出:存在无穷公式集S和公式A使得$S \Vdash A$，但对任意$S' \subsetneq_f S$，$S' \nVdash A$。我们下面证明之。令$S = \{B_i \in \mathscr{L}: i = 1, 2, \cdots\}$，其中$B_i =$ "$a_i \vee a_{i+1} \to a_{i+1} \wedge \neg a_i$"，其中这样的$a_i$是逻辑不相容的原子公式。令$A =$ "$a_1 \to \bot$"，从概率公理易得，对所有$i = 1, 2, \cdots$，若$P(B_i) > 2/3$，则

$P(a_i) \leqslant \frac{1}{2} P(a_{i+1})$. 证明如下:

设 $P(B_i) > 2/3$, 即 $P(a_i \vee a_{i+1} \to a_{i+1} \wedge \neg a_i) > 2/3$, 若 $P(a_i \vee a_{i+1}) = 0$, 据定理5.1, $P(a_i) = 0$, 得证. 若

$$P(a_i \vee a_{i+1}) \neq 0,$$

则

$$P((a_i \vee a_{i+1}) \wedge (a_{i+1} \wedge \neg a_i)) > 2/3(P(a_i \vee a_{i+1})),$$

因此

$$P(a_{i+1} \wedge \neg a_i) > 2/3(P(a_i) + P(a_{i+1}) - 0),$$

因为 a_i 和 a_{i+1} 逻辑不相容. 而据定义 $P(a_{i+1}) \geqslant P(a_{i+1} \wedge \neg a_i)$, 所以 $3P(a_{i+1}) > 2P(a_i) + 2P(a_{i+1})$, 因而

$$\frac{1}{2} P(a_{i+1}) > P(a_i).$$

因为 $a_1, a_2, \cdots$ 逻辑不相容,所以

$$P\left(\bigvee_{i=1}^{n} a_i\right) = \sum_{i=1}^{n} P(a_i).$$

又因为 $2P(a_i) < P(a_{i+1})$,所以

$$P\left(\bigvee_{i=1}^{n} a_i\right) > P(a_1) + 2P(a_1) + 2 \cdot 2P(a_1) + \cdots + 2^{n-1}P(a_1) = (2^n - 1)P(a_1).$$

若 $P(a_1) \neq 0$, 则存在 j, k ($j \leqslant k$, $j, k \in N$) 使得 $P(a_1) \geqslant j/k$, 因此总能取一有穷的 n 使得

$$P\left(\bigvee_{i=1}^{n} a_i\right) > 1,$$

而这与概率函数的定义矛盾. 因此 $P(a_1) = 0$, 从而 $P(a_1 \to \bot) = 1$, 亦即 $S \Vdash A$, 这是因为从 S 的所有公式至少有

概率 2/3 能确保 A 有任意高的概率. 但另一方面，同样也可以证明，对任意$S'\subseteq_f S$ 和赋值 $P(a_1)>0$，据定义 5.6(1)，$P(a_1\to\bot)=0$ 是和赋值 S' 的所有公式任意高的概率是一致的，因此 $S'\Vdash\!\!\!\!/\ A$.

此外我们还可以据定义 5.6 证明下列规则对合理后承推理不成立：对 $S\subseteq\mathscr{L}, \varphi, \psi\in\mathscr{L}^*$，若 $S,\varphi\Vdash\psi$，则 $S\Vdash\varphi\to\psi$ (这是因为，令 $\psi\in S$，则 $S\Vdash\psi$，亦即 $S,\neg\psi\Vdash\psi$，但 $S\Vdash\!\!\!\!/\ \neg\psi\to\psi$)。这说明类似古典命题演算的演绎定理对合理后承概念不成立.

下面我们首先研究限于有穷前提集的合理后承关系. 我们将证明这样的合理后承关系等价于一些有直观意义的条件，并且将给出自然演绎系统的一推理规则集，使得从有穷前提集 S 推出结论 A 当且仅当 A 能从 S 根据这些规则是可演绎的. Adams 是在"概率后承"的关系下给出这些推理规则集的. 可以证明 A 是 S 的概率后承的充分条件是，A 是 S 的合理后承. 至于必要条件，即完全性证明，我们将在以后研究.

§2 概率后承

Adams 用定义的形式给出从公式集 S 推演"概率后承"的自然演绎系统的一个推理规则集.

定义 5.7 令 $S\subseteq\mathscr{L}$，S 的概率后承集是包含 S 作为子集的最小集合 S' 且对所有 $\varphi,\psi,\theta\in\mathscr{L}^*$，下列条件满足：

(1) 若 φ 重言等值 ψ 且 $\varphi\to\theta\in S'$，则 $\psi\to\theta\in S'$.

(2) 若 $\varphi\in S'$，则 $\top\to\varphi\in S'$；若 $\top\to\varphi\in S'$，则 $\varphi\in S'$.

(3) 若 $\varphi\models\psi$，则 $\varphi\to\psi\in S'$.

(4) 若 $\varphi\to\theta\in S'$，且 $\psi\to\theta\in S'$，则 $\varphi\vee\psi\to\theta\in S'$.

(5) 若 $\varphi \vee \psi \to \theta \in S'$,且 $\psi \to \neg\theta \in S'$,则 $\varphi \to \theta \in S'$.

(6) 若 $\varphi \to \psi \wedge \theta \in S'$, 则 $\varphi \to \theta \in S'$.

(7) 若 $\varphi \to \psi \in S'$,且 $\varphi \to \theta \in S'$,则 $\varphi \to \psi \wedge \theta \in S'$.

(8) 若 $\varphi \to \psi \in S'$,且 $\varphi \wedge \psi \to \theta \in S'$,则 $\varphi \to \theta \in S'$.

实际上，当考虑 $S \subseteq \mathscr{L}$ 的概率后承集时，我们可以把(1)—(8)限制在 $\mathscr{L}$ 中,即 $S \subseteq \mathscr{L}$ 的概率后承类就是根据关于 $\mathscr{L}$ 的(1)—(8)从 S 推出的公式类. 还要注意,概率后承是 Adams 系统的主要语法概念. 为了方便起见,下面我们用 $S \vdash_p A$ 表示 A 是 S 的概率后承. 显然,所有的概率后承都是重言后承. 但逆是否成立? 下面几节我们来考虑这个问题.

定理 5.3 令 $\theta, \varphi_1, \cdots, \varphi_n, \psi_1, \cdots, \psi_n \in \mathscr{L}^*$,

(1) 若 $\psi_1 \models \theta$, 则 $\varphi_1 \to \psi_1 \vdash_p \varphi_1 \to \theta$.

(2) $\varphi_1 \to \psi_1 \vdash_p \varphi_1 \to \varphi_1 \wedge \psi_1$, $\varphi_1 \to \varphi_1 \wedge \psi_1 \vdash_p \varphi_1 \to \psi_1$.

(3) $\varphi_1 \to \psi_1 \vdash_p \varphi_1 \vee \varphi_2 \to \neg(\varphi_1 \wedge \neg\psi_1)$.

(4) $\{\varphi_i \to \psi_i : i = 1, \cdots, n\} \vdash_p \varphi_1 \vee \cdots \vee \varphi_n \to \neg(\varphi_1 \wedge \neg\psi_1) \wedge \cdots \wedge \neg(\varphi_n \wedge \neg\psi_n)$.

(5) 若 $\varphi_1 \wedge \psi_1 \models \varphi_2 \wedge \psi_2$, 且 $\varphi_2 \wedge \neg\psi_2 \models \varphi_1 \wedge \neg\psi_1$, 则 $\varphi_1 \to \psi_1 \vdash_p \varphi_2 \to \psi_2$.

证明: 在此我们只证(5),其余的留给读者. 我们假定 $\varphi_1 \to \psi_1$, 根据定义 5.7(1),因为 φ_1 与 $(\varphi_1 \wedge \varphi_2) \vee (\varphi_1 \wedge \neg\varphi_2)$ 重言等价,所以

$(\varphi_1 \wedge \varphi_2) \vee (\varphi_1 \wedge \neg\varphi_2) \to \psi_1$.

假定 $\varphi_1 \wedge \psi_1 \models \varphi_2 \wedge \psi_2$, 则

$\varphi_1 \wedge \neg\varphi_2 \models \neg\psi_1$.

根据定义 5.7(3), 则有

$\varphi_1 \wedge \neg\varphi_2 \to \neg\psi_1$,

同理，$\varphi_1\wedge\varphi_2\wedge\psi_1\models\psi_2$，因此根据 5.7(3)，有

$$\varphi_1\wedge\varphi_2\wedge\psi_1\rightarrow\psi_2.$$

根据定义 5.7(8)有，$\varphi_1\wedge\varphi_2\rightarrow\psi_2$，假定

$$\varphi_2\wedge\neg\psi_2\models\varphi_1\wedge\neg\psi_1,$$

我们有 $\varphi_2\wedge\neg\varphi_1\models\psi_2$. 根据定义 5.7(3)，得

$$\varphi_2\wedge\neg\varphi_1\rightarrow\psi_2.$$

从而由定义 5.7(4)，得 $(\varphi_1\wedge\varphi_2)\vee(\varphi_2\wedge\neg\varphi_1)\rightarrow\psi_2$. 因为 φ_2 与$(\varphi_1\wedge\varphi_2)\vee(\varphi_2\wedge\neg\varphi_1)$ 重言等价，且据定义 5.7(1)，有

$$\varphi_2\rightarrow\psi_2.$$

定理 5.4 令 $S\subseteq\mathscr{L}, A\in\mathscr{L}$，若 $S\vdash_p A$，则 $S\Vdash A$.

证明：只须证明定义 5.7 的规则(1)—(8)保持合理后承性. 下面我们分别加以验证.

(1) 的验证是容易的：若 φ 重言等价于 ψ，则对任意概率函数 P, $P(\varphi)=P(\psi)$，因此 $P(\varphi\rightarrow\theta)=P(\psi\rightarrow\theta)$，所以 $\varphi\rightarrow\theta\Vdash\psi\rightarrow\theta$.

验证(2)：因为 φ 和 $\top\wedge\varphi$ 重言等价，所以对任意概率函数 P，

$$P(\varphi)=P(\top\wedge\varphi).$$

又因为 $P(\top)=1$，所以

$$P(\varphi)=P(\top\wedge\varphi)/P(\top)=P(\top\rightarrow\varphi),$$

因此 $\varphi\Vdash\top\rightarrow\varphi$ 且 $\top\rightarrow\varphi\Vdash\varphi$.

验证(3)：只须证 $\varphi\models\psi\Rightarrow P(\varphi\rightarrow\psi)=1$. 若 $\varphi\models\psi$，则 φ 与 $\varphi\wedge\psi$ 重言等价，所以 $P(\varphi)=P(\varphi\wedge\psi)$. 因此若 $P(\varphi)\neq 0$，则 $P(\varphi\rightarrow\psi)=P(\varphi\wedge\psi)/P(\varphi)=1$. 若 $P(\varphi)=0$，仍有 $P(\varphi\rightarrow\psi)=1$.

容易用下列不等式验证 (4) 成立，

$$\frac{P(\psi\vee\mu)}{P(\varphi\vee\theta)}\leqslant\frac{P(\psi)}{P(\varphi)}+\frac{P(\mu)}{P(\theta)},\qquad(5.1)$$

其中 P 是任意概率函数，φ，ψ，θ，μ 是 $\mathscr{L}$ 的使得 $P(\varphi)$ 和 $P(\theta)$ 皆正的任意公式. 我们先来证明上述不等式成立. 令

$$P(\psi \wedge \neg\mu) = a, P(\psi \wedge \mu) = b, P(\mu \wedge \neg\psi) = c,$$

$$P(\varphi \wedge \neg\theta) = d, P(\varphi \wedge \theta) = e, P(\theta \wedge \neg\varphi) = f,$$

则

$$\begin{aligned} P(\psi \vee \mu) &= P((\psi \wedge \neg\mu) \vee (\psi \wedge \mu) \vee (\mu \wedge \neg\psi)) \\ &= a + b + c, \end{aligned}$$

同理

$$P(\varphi \vee \theta) = d + e + f,$$

$$P(\psi) = P((\psi \wedge \neg\mu) \vee (\psi \wedge \mu)) = a + b,$$

$$P(\mu) = b + c, P(\varphi) = d + e, P(\theta) = e + f.$$

因此只须证

$$\frac{a+b+c}{d+e+f} \leqslant \frac{a+b}{d+e} + \frac{b+c}{e+f}, \tag{5.2}$$

其中所有的数皆非负,所有的分母皆正. 要证明 (5.2)，只须进行简单的代数运算,请读者自证.

现在我们要证,对任意 $\varepsilon > 0$，若

$$P(\varphi \to \theta) > 1 - \frac{1}{2}\varepsilon$$

且

$$P(\psi \to \theta) > 1 - \frac{1}{2}\varepsilon,$$

则 $P(\varphi \vee \psi \to \theta) > 1 - \varepsilon$. 下面分情形考虑.

情形 1：假设 $P(\varphi) > 0$ 且 $P(\psi) > 0$，则 $P(\varphi \vee \psi) > 0$，因此,根据定理 5.1 和 (5.1)，有

$$\begin{aligned} &P(\varphi \vee \psi \to \theta) \\ &= 1 - P(\varphi \vee \psi \to \neg\theta) \\ &= 1 - P((\varphi \vee \psi) \wedge \neg\theta)/P(\varphi \vee \psi) \end{aligned}$$

$$= 1 - P((\varphi \wedge \neg\theta) \vee (\psi \wedge \neg\theta))/P(\varphi \vee \psi)$$
$$\geq 1 - P(\varphi \wedge \neg\theta)/P(\varphi) - P(\psi \wedge \neg\theta)/P(\psi)$$
$$= 1 - P(\varphi \to \neg\theta) - P(\psi \to \neg\theta)$$
$$= P(\varphi \to \theta) + P(\psi \to \theta) - 1$$
$$> 1 - \frac{1}{2}\varepsilon + 1 - \frac{1}{2}\varepsilon - 1$$
$$= 1 - \varepsilon.$$

情形 2：假设 $P(\varphi)$ 和 $P(\psi)$ 中有一为 0．设 $P(\varphi)=0$，$P(\psi) \neq 0$，因为$\psi \to \theta$ 是 $\varphi \vee \psi \to \theta$ 的重言后承，所以 $P(\varphi \vee \psi \to \theta) \leqslant P(\psi \to \theta)$．另一方面，我们来证明

$$P(\varphi \vee \psi \to \theta) \geqslant P(\psi \to \theta),$$

假设 $P(\varphi \vee \theta) \neq 0$，否则 $P(\varphi \vee \psi \to \theta) = 1$，显然不等式成立．因此根据定理 5.1

$$P(\varphi \vee \psi \to \theta)$$
$$= P((\varphi \vee \psi) \wedge \theta)/P(\varphi \vee \psi)$$
$$= P((\varphi \vee \psi) \wedge \theta)/(P(\varphi) + P(\psi) - P(\varphi \wedge \psi)).$$

因为 $P(\varphi) = 0$，所以 $P(\varphi \wedge \psi) = 0$，因此上式变为

$$= P((\varphi \vee \psi) \wedge \theta)/P(\psi).$$

又因为 $(\varphi \vee \psi) \wedge \theta$ 是 $\psi \wedge \theta$ 的重言后承，所以上式变为

$$\geqslant P(\psi \wedge \theta)/P(\psi) = P(\psi \to \theta).$$

情形 3：假设 $P(\varphi) = P(\psi) = 0$，据定理 5.1，

$$P(\varphi \vee \psi) = 0,$$

则 $P(\varphi \vee \psi \to \theta) = 1$．

验证 (5) 只须证

$$P(\varphi \vee \psi \to \theta) + P(\psi \to \neg\theta) \leqslant 1 + P(\varphi \to \theta).$$

假设 $P(\varphi) \neq 0$ (否则显然)．

(i) 若 $P(\psi) = 0$，则 $P(\psi \to \theta) = 1$，且

$$P(\varphi \vee \psi \to \theta) = P((\varphi \wedge \theta) \vee (\psi \wedge \theta))/P(\varphi \vee \psi)$$
$$= (P(\varphi \wedge \theta) + P(\psi \wedge \theta) - P(\varphi \wedge \theta \wedge \psi \wedge \theta))/$$
$$(P(\varphi) + P(\psi) - P(\varphi \wedge \psi))$$
$$= P(\varphi \wedge \theta)/P(\varphi) = P(\varphi \to \theta).$$

(ii) 若 $P(\psi) \neq 0$，据不等式 (5.1)

$$\frac{P((\varphi \wedge \theta) \vee (\psi \wedge \theta))}{P(\varphi \vee \psi)} \leqslant \frac{P(\varphi \wedge \theta)}{P(\varphi)} + \frac{P(\psi \wedge \theta)}{P(\psi)},$$

即

$$P(\varphi \vee \psi \to \theta) \leqslant P(\varphi \to \theta) + P(\psi \to \theta),$$
$$P(\varphi \vee \psi \to \theta) + P(\psi \to \neg\theta) \leqslant 1 + P(\varphi \to \theta),$$

因此当 $P(\varphi \vee \psi \to \theta)$ 和 $P(\psi \to \neg\theta)$ 都大于 $1 - \frac{1}{2}\varepsilon$ 时，

$P(\varphi \to \theta) > 1 - \varepsilon$.

(6) 的验证只须证明

$$P(\varphi \to \psi \wedge \theta) \leqslant P(\varphi \to \theta).$$

假设 $P(\varphi) \neq 0$ (否则显然). 因为 $\varphi \wedge (\psi \wedge \theta) \models \varphi \wedge \psi$，所以 $P(\varphi \wedge (\psi \wedge \theta)) \leqslant P(\varphi \wedge \psi)$，因此 $P(\varphi \wedge (\psi \wedge \theta))/P(\varphi) \leqslant P(\varphi \wedge \psi)/P(\varphi)$，即 $P(\varphi \to \psi \wedge \theta) \leqslant P(\varphi \to \theta)$.

(7) 的验证只须证明

$$P(\varphi \to \psi) + P(\varphi \to \theta) \leqslant 1 + P(\varphi \to \psi \wedge \theta).$$

假设 $P(\varphi) \neq 0$ (否则显然). 据定理 5.1，我们有

$$P(\varphi \wedge \psi) + P(\varphi \wedge \theta) - P((\varphi \wedge \psi) \vee (\varphi \wedge \theta))$$
$$= P(\varphi \wedge (\psi \wedge \theta)),$$

所以

$$P(\varphi \to \psi) + P(\varphi \to \theta)$$
$$= P(\varphi \to \psi \wedge \theta) + P((\varphi \wedge \psi) \vee (\varphi \wedge \theta))/P(\varphi),$$

因此

$$P(\varphi \to \psi) + P(\varphi \to \theta) \leqslant P(\varphi \to \psi \wedge \theta) + 1.$$

(8) 设 $P(\varphi)\neq 0$（否则显然）. 若 $P(\varphi\wedge\psi)=0$，则 $P(\varphi\wedge\psi\to\theta)=1, P(\varphi\to\psi)=0$，因此总有

$$P(\varphi\to\psi)+P(\varphi\wedge\psi\to\theta)\leqslant 1+P(\varphi\to\theta),$$

因此若

$$P(\varphi\to\psi)>1-\frac{1}{2}\varepsilon$$

且

$$P(\varphi\wedge\psi\to\theta)>1-\frac{1}{2}\varepsilon,$$

则 $P(\varphi\to\theta)>1-\varepsilon$.

若 $P(\varphi\wedge\psi)\neq 0$，则据假设，有

$$\begin{aligned}&P(\varphi\to\psi)\cdot P(\varphi\wedge\psi\to\theta)\\&=(P(\varphi\wedge\psi)/P(\varphi))\\&\quad\cdot(P(\varphi\wedge\psi\wedge\theta)/P(\varphi\wedge\psi))\\&=P(\varphi\wedge\psi\wedge\theta)/P(\varphi)\\&>\left(1-\frac{1}{2}\varepsilon\right)^2,\end{aligned}$$

因此

$$\begin{aligned}P(\varphi\to\theta)&=P(\varphi\wedge\theta)/P(\varphi)\\&\geqslant P(\varphi\wedge\psi\wedge\theta)/P(\varphi)\\&>1-\varepsilon+\varepsilon^2>1-\varepsilon.\end{aligned}$$

注意：(1) 定理 5.4 可以看作是 Adams 系统中的可靠性定理. (2) 据定理 5.2(2) 证明中的说明，有些重言后承不是合理后承，所以我们已知有些重言后承不是概率后承.

§3 P 偏序以及与之关联的齐一概率序列

在上一节中，我们已证明：$S\vdash_p A\Rightarrow S\Vdash A$. 这一节和下一节要证明逆也成立. 为此要系统地考虑概率函数和下列概

率函数序列的构造：S 的公式的概率趋于 1 作为极限，而 A 的概率的极限小于 1．构造这样的概率函数序列的一个相当方便的方式是引入 P 偏序概念．直观上说，P 偏序是 $\mathscr{L}^*$ 上的某种弱偏序关系，这一关系可以作如下解释：若 φ 和 ψ 是真值函数公式，$\leqslant$ 是一 P 偏序且 $\varphi \leqslant \psi$ 成立，则 $P(\psi)/P(\varphi)$ 不趋于 0，或者若严格不等式 $\varphi < \psi$ 成立，则 $P(\varphi)/P(\psi)$ 趋于 0．我们首先形式地定义"P 偏序"，然后再说明此定义与前面的直观解释相符．

定义 5.8 令 $\mathscr{L}$ 是一语言，我们有

(1) $\mathscr{L}$ 的 P 偏序 $\leqslant$ 是 $\mathscr{L}^*$ 上的二元关系且满足下列性质：对所有 $\varphi,\psi,\theta \in \mathscr{L}^*$，有

(i) $\varphi \leqslant \psi$ 或 $\psi \leqslant \varphi$，且若 $\varphi \leqslant \psi$，$\psi \leqslant \theta$，则 $\varphi \leqslant \theta$（即 $\leqslant$ 是弱偏序）．

(ii) 若 $\varphi \models \psi$，则 $\varphi \leqslant \psi$，且 $\bot < \top$（即并非 $\top \leqslant \bot$）．

(iii) $\varphi \vee \psi \leqslant \theta \Longleftrightarrow \varphi \leqslant \theta$ 且 $\psi \leqslant \theta$．

(iv) $\varphi \leqslant \psi \vee \theta \Longleftrightarrow \varphi \leqslant \psi$ 或 $\varphi \leqslant \theta$．

（注意 (iii) 和 (iv) 的一个方向可从 (ii) 得到）．

(2) 若 $\leqslant$ 是 $\mathscr{L}$ 的 P 偏序，φ，$\psi \in \mathscr{L}^*$，则 φ 在 $\leqslant$ 中成立 $\Longleftrightarrow \neg\varphi < \varphi$（即并非 $\varphi \leqslant \neg\varphi$）；$\varphi \rightarrow \psi$ 在 $\leqslant$ 中成立 $\Longleftrightarrow$ $\varphi \wedge \neg\psi < \varphi \wedge \psi$ 或 $\varphi \leqslant \bot$．

φ 或 $\varphi \rightarrow \psi$ 在 P 偏序 $\leqslant$ 成立的直观意义是 $P(\varphi)$ 或 $P(\varphi \rightarrow \psi)$ 的极限为 1．

定理 5.5 令 $\mathscr{L}$ 是有穷语言，$\leqslant$ 是 $\mathscr{L}^*$ 上的二元关系，则 $\leqslant$ 是 P 偏序 $\Longleftrightarrow$ 存在 $\mathrm{SD}\mathscr{L}^+ = \mathrm{SD}\mathscr{L} \cup \{\bot\}$ 上的弱偏序 $\leqslant_0$ 使得

$$\begin{cases} \bot \leqslant_0 \alpha, \text{ 对所有的 } \alpha \in \mathrm{SD}\mathscr{L}^+, \\ \bot <_0 \alpha, \text{ 对某个 } \alpha \in \mathrm{SD}\mathscr{L}^+. \end{cases} \tag{5.3}$$

对所有 $\varphi,\psi \in \mathscr{L}^*$，若 $\alpha,\beta \in \mathrm{SD}\mathscr{L}^+$ 且分别极大蕴涵 φ,ψ

即对所有的 $\alpha',\beta' \in SD\mathscr{L}^+$，其中 $\alpha' \models \varphi, \beta' \models \psi$，有 $\alpha' \leqslant_0 \alpha$，$\beta' \leqslant_0 \beta$，也称 α, β 分别是 φ，ψ 在 $SD\mathscr{L}^+$ 中的极大蕴涵(元)，则

$$\varphi \leqslant \psi \Longleftrightarrow \alpha \leqslant_0 \beta. \qquad (5.4)$$

若$\leqslant$和$\leqslant_0$满足上述条件 (5.3) 和 (5.4)，则称 $\leqslant$ 是由 $\leqslant_0$ 确定的 P 偏序。

证明："$\Rightarrow$"．假设$\leqslant$是 $\mathscr{L}^*$ 的P 偏 序，令 $\leqslant_0 = \leqslant \upharpoonright SD\mathscr{L}^+$，只须证$\leqslant$和$\leqslant_0$满足(5.3)和 (5.4)．显然$\leqslant_0$是 $SD\mathscr{L}^+$ 上的弱偏序，因为$\leqslant$是弱偏序，且弱偏序对一子域限制仍是该子域上的弱偏序．对所有 $\alpha \in SD\mathscr{L}^+$，$\perp \leqslant_0 \alpha$ 也同样可以从定义 5.8(1) 的 (ii) 直接得到．对某个 $\alpha \in SD\mathscr{L}^+$，$\perp <_0 \alpha$ 可以从定义 5.8(1) 的 (iii) 得到，因为若对任意 $\alpha \in SD\mathscr{L}^+$，并非$\perp <_0 \alpha$，则对所有 $\alpha \in SD\mathscr{L}^+$，$\alpha \leqslant \perp$成立．据定义 5.8 (1) 的 (iii) 和 $SD\mathscr{L}^+$ 的性质，$U \leqslant \perp$成立(其中U是所有 $\alpha \in SD\mathscr{L}^+$ 的析取)．但是 $SD\mathscr{L}$ 的所有元素的析取(在这个语言中)重言等价$\top$，因此$\top \leqslant \perp$，得出矛盾．

为了证明 (5.4) 成立，我们先假设 α，$\beta \in SD\mathscr{L}^+$ 分 别在$\leqslant$下极大蕴涵 φ，ψ，此外还能不失一般性地假设φ和ψ是 $SD\mathscr{L}^+$ 中某些元素的析取，因为根据状态描述的构造，φ和ψ重言等价于 $SD\mathscr{L}^+$ 的某些元素的析取．从定义 5.8(1) 的 (ii) 可直接得到，重言等价的公式可以在$\leqslant$中互相替换．在这样的情况下，易证一状态描述重言蕴涵这样一析取当且仅当它是后者的一个析取肢．

下面我们来证明 $\varphi \leqslant \psi \Longleftrightarrow \alpha \leqslant \beta$． 为此先证明 $\varphi \leqslant \psi \Longleftrightarrow \alpha \leqslant \psi$，然后再证明 $\alpha \leqslant \psi \Longleftrightarrow \alpha \leqslant \beta$．根据定义 5.8 (1) 的 (i) 和 (iii)，我们有 $\varphi \leqslant \psi \Longleftrightarrow \alpha \leqslant \psi$．由于(iii)，和 (i) 的传递性，对 $SD\mathscr{L}^+$ 中所有的重言蕴涵φ的α'，有

$$\alpha' \leqslant \psi \Longleftrightarrow \alpha \leqslant \psi \Longleftrightarrow \varphi \leqslant \psi,$$

其中 $\alpha \in \mathrm{SD}\mathscr{L}^+$ 且极大蕴涵 φ。据 (i) 和 (iv) 有

$$\alpha \leqslant \psi \Longleftrightarrow \alpha \leqslant \beta,$$

这是因为已假设 ψ 是 $\mathrm{SD}\mathscr{L}^+$ 的某些元素的析取，所以从 (iv) 可得，对 ψ 中某个析取肢 β'，$\alpha \leqslant \psi \Longleftrightarrow \alpha \leqslant \beta'$ (再据$\leqslant$的传递性推出) $\Longleftrightarrow \alpha \leqslant \beta$，其中 $\beta \in \mathrm{SD}\mathscr{L}^+$ 且极大蕴涵 ψ，这样就证明了 (5.4)。因此若$\leqslant$是 $\mathscr{L}^*$ 的 P 偏序，则存在 $\mathrm{SD}\mathscr{L}^+$ 上满足 (5.3) 和 (5.4) 的弱偏序 $\leqslant_0$。下面证"$\Leftarrow$"。

假设$\leqslant_0$是 $\mathrm{SD}\mathscr{L}^+$ 上的弱偏序且 $\leqslant_0$ 和$\leqslant$满足 (5.3) 和 (5.4)，要证这样的$\leqslant$满足定义 5.8(1)的(i)—(iv)。假设 $\alpha, \beta, \gamma \in \mathrm{SD}\mathscr{L}^+$ 且分别极大蕴涵 φ, ψ，$\theta \in \mathscr{L}^*$，据 (5.4)，$\varphi \leqslant \psi \Longleftrightarrow \alpha \leqslant_0 \beta$。因为 $\leqslant_0$ 是弱偏序，所以 $\alpha \leqslant_0 \beta$ 或 $\beta \leqslant_0 \alpha$，因此 $\varphi \leqslant \psi$ 或 $\psi \leqslant \varphi$。又因为$\leqslant_0$具有传递性，所以$\leqslant$在 $\mathscr{L}^*$ 上也有传递性，因此$\leqslant$满足 (i)。

假设 $\varphi \models \psi$，则存在 φ', ψ' 使得它们是 $\mathrm{SD}\mathscr{L}^+$ 中某些元素的析取且分别重言等价于 φ，ψ，因此 $\varphi' \models \psi'$。显然 φ' 是 ψ' 的子析取，否则 φ' 中存在一析取肢不是 ψ' 的析取肢，这时若对这个析取肢赋值真，而对 ψ' 中所有的析取肢赋值假，则在这个赋值下 φ' 真而 ψ' 假，与 $\varphi' \models \psi'$ 矛盾。根据 $\mathrm{SD}\mathscr{L}^+$ 的性质和$\leqslant_0$的性质以及 φ'，ψ' 的构造，存在 α，$\beta \in \mathrm{SD}\mathscr{L}^+$ 使得它们分别极大蕴涵 φ', ψ'，且 α，β 分别是 φ'，ψ' 的析取肢，因为它们的析取肢就是 $\mathrm{SD}\mathscr{L}^+$ 的元素且重言蕴涵 φ', ψ'，而极大性是根据$\leqslant_0$的性质。

现证 α, β 必然是 φ'，ψ' 的析取肢，否则，对 α，β 赋值真(若 φ', ψ' 不重言等价，则 α, β 也不重言等价，否则与极大性矛盾)，对 φ' 和 ψ' 中所有的析取肢赋值假，这样的赋值与重言蕴涵性矛盾。因为 φ' 是 ψ' 的子析取，所以 α 也在 ψ' 中，因此 $\alpha \leqslant_0 \beta$，据 (5.4)，$\varphi \leqslant \psi$。

据 (5.3)，存在 $\alpha \in \mathrm{SD}\mathscr{L}^+$ 使得 $\perp <_0 \alpha$，显然 $\alpha \models \top$，

所以在 $SD\mathscr{L}^+$ 中必有⊤的极大蕴涵元 α^*: $\alpha \leqslant_0 \alpha^*$，从而 $\bot <_0 \alpha^*$. 据 $SD\mathscr{L}^+$ 的构造，⊥的极大蕴涵元只能是自身(因为任意 $\alpha \in SD\mathscr{L}$ 不是永假公式，所以 $\alpha \nvDash \bot$)，因此据 (5.4)，$\bot < \top$，亦即≤满足 (ii).

据 (5.3) 和 (5.4)，(iii) 的"⇒"显然，现证"⇐". 假设 α，$\beta, \gamma \in SD\mathscr{L}^+$ 且分别极大蕴涵 $\varphi, \psi, \theta \in \mathscr{L}^*$，据 (5.4)，$\varphi \leqslant \theta$ 且 $\psi \leqslant \theta \Longleftrightarrow \alpha \leqslant_0 \gamma$ 且 $\beta \leqslant_0 \gamma$. 考虑 $\varphi \vee \psi$，令 $(\varphi \vee \psi)'$ 是 $SD\mathscr{L}^+$ 中某些元素的析取且与 $\varphi \vee \psi$ 重言等价，显然 α, β 皆是 $(\varphi \vee \psi)'$ 的析取肢. 根据 $(\varphi \vee \psi)'$ 的构造和 α, β 的性质，其中必有 $\varphi \vee \psi$ 的极大蕴涵元，即若 $\alpha \leqslant \beta$，β 就是 $(\varphi \vee \psi)'$ 的极大蕴涵元，否则据 $\leqslant_0$ 的性质，α 是 $\varphi \vee \psi$ 的极大蕴涵元. 据题设和 (5.4)，$\varphi \vee \psi \leqslant \theta$. 亦即≤满足 (iii).

最后证 ≤ 满足 (iv). $\varphi \leqslant \psi$ 或 $\varphi \leqslant \theta \Longleftrightarrow$ 存在 α, β，$\gamma \in SD\mathscr{L}^+$ 使得它们分别是 φ, ψ, θ 的极大蕴涵元，且据(5.4)，$\alpha \leqslant_0 \beta$ 或 $\alpha \leqslant_0 \gamma$，据 $(\psi \vee \theta)'$ 的构造和 β, γ 的性质，β 或 γ 就是 $(\psi \vee \theta)'$ 的极大蕴涵元，因此据 (5.4)，$\varphi \leqslant \psi \vee \theta$. 证毕.

现在我们考虑 P 偏序与齐一概率序列的联系，从这种联系我们可以把 $\varphi \leqslant \psi$ 解释为 $P(\psi)/P(\varphi)$ 的极限为正. 这个思想表现在以下定义中.

定义 5.9 令≤是 $\mathscr{L}^*$ 上的 P 偏序，$P_1, P_2, \cdots$ 是 $\mathscr{L}$ 的齐一概率函数序列，称 $P_1, P_2, \cdots$ 是与≤关联的齐一概率函数序列，若对所有 $\varphi, \psi \in \mathscr{L}^*$，

$$\varphi \leqslant \psi \Longleftrightarrow \lim_{n \to \infty} (P_n(\varphi \vee \psi \to \psi)) > 0.$$

下面定理说明：只要极限存在，则

$$\lim_{n \to \infty} (P_n(\varphi \vee \psi \to \psi)) > 0 \Longleftrightarrow \lim_{n \to \infty} (P_n(\psi)/P_n(\varphi)) > 0.$$

因此 $\varphi \leqslant \psi$ 的直观解释是正当的. 同样，从定义 5.9 很容易

得到，若 $P_1,P_2,\cdots$ 与≤关联，则

$$\varphi < \psi \Leftrightarrow \lim_{n\to\infty}(P_n(\varphi \vee \psi \to \varphi)) = 0,$$

而这又等价于(若该极限存在)：

$$\lim_{n\to\infty}(P_n(\varphi)/P_n(\psi)) = 0.$$

定理 5.6 令 $\mathscr{L}$ 有穷，$A \in \mathscr{L}$，$S \subseteq_f \mathscr{L}$，则

(1) 若 $P_1, P_2, \cdots$ 是 $\mathscr{L}$ 的齐一概率函数序列，则存在 $\mathscr{L}$ 的唯一的 P 偏序≤使得 $P_1,P_2,\cdots$ 与≤关联。

(2) 若≤是 $\mathscr{L}$ 的 P 偏序，则存在 $\mathscr{L}$ 的与≤关联的齐一概率函数序列。

(3) 若≤是 $\mathscr{L}$ 的 P 偏序且 $P_1, P_2, \cdots$ 是与≤关联的齐一序列，则 A 在≤中成立 $\Leftrightarrow \lim_{n\to\infty} P_n(A) = 1$.

(4) 若 $S \Vdash A$，则 A 在所有使 S 的全部公式成立 P 偏的序中成立。

证明：(1) 只须证由下列条件：对所有 $\varphi,\psi \in \mathscr{L}^*$，

$$\varphi \leqslant \psi \Leftrightarrow \lim_{n\to\infty}(P_n(\varphi \vee \psi \to \psi)) > 0$$

定义的二元关系≤满足定义 5.8 (1) 的 (i)—(iv)。另外据定义 5.9 可证，$P_1,P_2,\cdots$ 是与≤关联的齐一序列。

显然对任意 $\varphi,\psi \in \mathscr{L}^*$，

$$\lim_{n\to\infty}(P_n(\varphi \vee \psi \to \varphi \vee \psi)) = 1$$

$$\leqslant \lim_{n\to\infty} P_n(\varphi \vee \psi \to \varphi) + \lim_{n\to\infty} P_n(\varphi \vee \psi \to \psi).$$

因为

$$\varphi \vee \psi \models ((\varphi \vee \psi) \wedge \varphi) \vee ((\varphi \vee \psi) \wedge \psi),$$

所以

$$P(\varphi \vee \psi) \leqslant P(((\varphi \vee \psi) \wedge \varphi) \vee ((\varphi \vee \psi) \wedge \psi)).$$

若 $P(\varphi \vee \psi) \neq 0$ (否则，显然)，则

$$1 \leqslant P(((\varphi \vee \psi) \wedge \varphi) \vee ((\varphi \vee \psi) \wedge \psi))/P(\varphi \vee \psi),$$

据定理 5.4 的证明中的不等式（5.1）等，

$$\frac{P(((\varphi\vee\psi)\wedge\varphi)\vee((\varphi\vee\psi)\wedge\psi))}{P(\varphi\vee\psi)}$$

$$\leqslant\frac{P((\varphi\vee\psi)\wedge\varphi)}{P(\varphi\vee\psi)}+\frac{P((\varphi\vee\psi)\wedge\psi)}{P(\varphi\vee\psi)},$$

所以

$$P(\varphi\vee\psi\rightarrow\varphi\vee\psi)\leqslant P(\varphi\vee\psi\rightarrow\varphi)+P(\varphi\vee\psi\rightarrow\psi),$$

即

$$1\leqslant\lim_{n\to\infty}P_n(\varphi\vee\psi\rightarrow\varphi)+\lim_{n\to\infty}P_n(\varphi\vee\psi\rightarrow\psi).$$

我们可得

$\lim\limits_{n\to\infty}(P_n(\varphi\vee\psi\rightarrow\varphi))>0$ 或 $\lim\limits_{n\to\infty}(P_n(\varphi\vee\psi\rightarrow\psi))>0.$

据定义 5.9，$\varphi\leqslant\psi$ 或 $\psi\leqslant\varphi$，所以二元关系$\leqslant$满足定义 5.8（1）的（i）。

二元关系$\leqslant$的传递性可以从定理 5.1 的（7）得到：对任意概率函数 P_n，

$$P_n(\varphi\vee\theta\rightarrow\theta)\geqslant P_n(\varphi\vee\psi\rightarrow\psi)\cdot P_n(\psi\vee\theta\rightarrow\theta).$$

假设 $\varphi\leqslant\psi$ 且 $\psi\leqslant\theta$，则

$$\lim_{n\to\infty}P_n(\varphi\vee\psi\rightarrow\psi)>0$$

且

$$\lim_{n\to\infty}P_n(\psi\vee\theta\rightarrow\theta)>0,$$

所以

$$\lim_{n\to\infty}P_n(\varphi\vee\theta\rightarrow\theta)>0,$$

因此 $\varphi\leqslant\theta$.

对于（ii）。①假设 $\varphi\models\psi$，所以 $\varphi\vee\psi\models(\varphi\vee\psi)\wedge\psi$，又因为 $(\varphi\vee\psi)\wedge\psi\models\varphi\vee\psi$，所以

$$P_n((\varphi\vee\psi)\wedge\psi)=P_n(\varphi\vee\psi).$$

设 $P_n(\varphi\vee\psi)\neq0$（否则易证），则

$$\lim_{n\to\infty} P_n(\varphi \vee \psi \to \psi)$$
$$= \lim_{n\to\infty} (P_n((\varphi \vee \psi) \wedge \psi)/P_n(\varphi \vee \psi))$$
$$= 1 > 0,$$
所以 $\varphi \leqslant \psi$. ②因为
$$P_n(\top \vee \bot) = P_n(\top) + P_n(\bot) - P_n(\top \wedge \bot)$$
$$= P_n(\top) = 1 > 0,$$
所以
$$\lim_{n\to\infty} P_n(\top \vee \bot \to \bot) = \lim_{n\to\infty} (P_n((\top \vee \bot) \wedge \bot)/ P_n(\top \vee \bot)) = 0,$$
即 $\bot < \top$.

对于 (iii). (iii) 的"⇒"可以从 (ii) 和 (i) 得到. 这是因为 $\varphi \models \varphi \vee \psi$, 所以据 (ii), $\varphi \leqslant \varphi \vee \psi$, 因此若 $\varphi \vee \psi \leqslant \theta$, 则据 (i) 的≤的传递性, $\varphi \leqslant \theta$. 同理可证 $\psi \leqslant \theta$. 现在我们来考虑"⇐". 只须证
$$P_n(\varphi \vee \theta \to \theta) > 0 \text{ 且 } P_n(\psi \vee \theta \to \theta) > 0$$
$$\Rightarrow P_n(\varphi \vee \psi \vee \theta \to \theta) > 0,$$
而这就是定理 5.1 的 (8).

对于 (iv). (iv) 的"⇐"能据 (i) 和 (ii) 得到. "⇒"能据定理 5.1 的(9)得到. 详细证明留给读者. 证毕.

证 (2) 令≤是 $\mathscr{L}^*$ 上的 P 偏序, $\mathrm{SD}\mathscr{L}^+ = \mathrm{SD}\mathscr{L} \cup \{\bot\}$, 假设 $\mathrm{SD}\mathscr{L}^+$ 中的元素以下列方式排序 (≈表示等价关系)
$$\bot \approx \alpha_1 \approx \cdots \approx \alpha_{n_1} < \alpha_{n_1+1} \approx \cdots \approx \alpha_{n_2} < \cdots < \alpha_{n_{K-1}+1} \approx \cdots \approx \alpha_{n_K},$$
即 $\mathrm{SD}\mathscr{L}^+$ 的元素根据等价类排序: $\alpha_i < \alpha_j \Leftrightarrow \alpha_i/\approx$ 在 $\alpha_j/\approx$ 左边(这里 $\alpha_i/\approx$ 表示以 α_i 为代表元的相对≈的等价

类，下同）。对每一 i，令 β_i 是第 i 个等价类中的元素的析取，即

$$\beta_i = \alpha_{n_{i-1}+1} \vee \cdots \vee \alpha_{n_i}.$$

我们首先对 $n = 2,3,\cdots$ 和 $i = 1,\cdots,K$，赋值 $P_n(\beta_i)$ 的值。若 $K = 2$，证明就很容易，对所有 $n = 2,3,\cdots$，我们令 $P_n(\beta_1) = 0, P_n(\beta_2) = 1$。若 $K > 2$，概率 P_n 定义如下：对 $n = 2,3,\cdots$，

$$P_n(\beta_1) = 0,$$

$$P_n(\beta_i) = (n^{2K-i})^{-1}，\text{对 } i = 2,3,\cdots,K-1,$$

$$P_n(\beta_K) = 1 - \frac{1}{n^{2K-2}} \cdot \frac{n^{K-2}-1}{n-1},$$

简单说，离⊥越远的 β_i 得到的概率越大。用代数变换，我们可得

$$\sum_{i=1}^{K} P_n(\beta_i) = 1，\text{对 } n = 2,3,\cdots, \tag{5.5}$$

$$P_n(\beta_i)/P_n(\beta_i+1) \leqslant 1/n,$$

$$\text{对 } i = 1,\cdots,K-1, n = 2,3,\cdots. \tag{5.6}$$

$SD\mathscr{L}^+$ 的单个元素的概率定义如下：对 $j = 1,\cdots,n_K$，若 α_j 是 β_i 的一析取肢，则

$$P_n(\alpha_j) = P(\beta_i)/\gamma_i,$$

其中 γ_i 是 β_i 的析取肢的个数。

显然概率 P_n 形成 $SD\mathscr{L}$ 上一分布，因而对语言 $\mathscr{L}$ 定义了唯一的概率函数 P_n，而且 $P_1, P_2, \cdots$ 形成 $\mathscr{L}$ 的齐一概率函数序列，现在剩下的只是证明该序列是与≤相关的齐一序列。

据定理 5.6(1)，只须证对任意 $\alpha_h, \alpha_j \in SD\mathscr{L}^+$，

$$\alpha_h \leqslant \alpha_j \Longleftrightarrow \lim_{n\to\infty}(P_n(\alpha_h \vee \alpha_j \to \alpha_j)) > 0,$$

因为据定理 5.6(1)，序列 $P_2, P_3, \cdots$ 与 $\mathscr{L}$ 的某个唯一的 P 偏序≤相关联，且据定理 5.5，这个 P 偏序≤自身又由它对

$SD\mathscr{L}^+$ 的限制$\leqslant_0$唯一确定。 至此完成本证明就比较容易.

若 $\alpha_h \leqslant \alpha_j$, 则 $\alpha_h \approx \alpha_j$ 或 $\alpha_h < \alpha_j$.

在第一种情况下, $P_n(\alpha_h) = P_n(\alpha_j)$, 由此可得 $P_n(\alpha_h \vee \alpha_j \to \alpha_j) \geqslant 1/2$, 因为若 $P_n(\alpha_h \vee \alpha_j) = 0$, 则显然

$$P_n(\alpha_h \vee \alpha_j \to \alpha_j) = 1 \geqslant 1/2.$$

若

$$P_n(\alpha_h \vee \alpha_j) \neq 0,$$

则

$$\begin{aligned} P_n(\alpha_h \vee \alpha_j \to \alpha_j) &= P_n((\alpha_h \vee \alpha_j) \wedge \alpha_j)/P_n(\alpha_h \vee \alpha_j) \\ &= (P_n(\alpha_h \vee \alpha_j) + P(\alpha_j) - P_n(\alpha_h \vee \alpha_j \vee \alpha_j))/P(\alpha_h \vee \alpha_j) \\ &= P_n(\alpha_j)/(P_n(\alpha_h) + P_n(\alpha_j) - P_n(\alpha_h \wedge \alpha_j)) \\ &= P_n(\alpha_j)/(P_n(\alpha_h) + P_n(\alpha_j)) = \frac{1}{2}. \end{aligned}$$

因此

$$\lim_{n\to\infty} P_n(\alpha_h \vee \alpha_j \to \alpha_j) \geqslant \frac{1}{2} > 0.$$

若 $\alpha_h<\alpha_j$, 则 α_h 和 α_j 分别是 β_p 和 $\beta_q(p<q)$ 的析取肢,此时,

$$P_n(\beta_p)/P_n(\beta_q) = \frac{1}{\gamma_p}\ P_n(\alpha_h) \Big/ \frac{1}{\gamma_q}\ P_n(\alpha_j) \leqslant \frac{1}{n},$$

其中 γ_p 和 γ_q 分别是 β_p 和 β_q 的析取肢的个数。因此

$$P_n(\alpha_h)/P_n(\alpha_j) \leqslant \frac{1}{n}\ (\gamma_q/\gamma_p).$$

显然

$$\lim_{n\to\infty} (P_n(\alpha_h)/P_n(\alpha_j)) = 0,$$

所以

$$\begin{aligned} \lim_{n\to\infty} P_n(\alpha_h \vee \alpha_j \to \alpha_j) &= \lim_{n\to\infty} (P_n(\alpha_j)/P_n(\alpha_h \vee \alpha_j)) \\ &= \lim_{n\to\infty} (P_n(\alpha_j)/(P_n(\alpha_j) + P_n(\alpha_h))) \\ &= \lim_{n\to\infty} \frac{1}{1 + \dfrac{P_n(\alpha_h)}{P_n(\alpha_j)}} = 1 > 0. \end{aligned}$$

证(3) 假设≤是P偏序，$P_1, P_2, \cdots$是与≤关联的齐一序列。若A不是条件句(例如$A=\varphi$)，据定义5.8(2)，A在≤中成立$\Leftrightarrow \neg\varphi < \varphi$，因为据定义5.9，$P_1, P_2, \cdots$与≤关联$\Leftrightarrow \lim\limits_{n\to\infty} P_n(\varphi \vee \neg\varphi \to \neg\varphi) = 0 \Leftrightarrow \lim\limits_{n\to\infty} P_n(\varphi) = 1$，这是因为$P_n(\varphi \vee \neg\varphi \to \neg\varphi) = 1 - P_n(\varphi)/P_n(\varphi \vee \neg\varphi) = 1 - P_n(\varphi)$。若$A$是条件句(不妨令$A = \varphi \to \psi$)，据定义5.8(2)，$A$在≤中成立$\Leftrightarrow \varphi \wedge \neg\psi < \varphi \wedge \psi$或$\varphi \leqslant \bot$。

情况1．据定义5.9，$\varphi \wedge \neg\psi < \varphi \wedge \psi \Leftrightarrow$并非$(\varphi \wedge \psi \leqslant \varphi \wedge \neg\psi) \Leftrightarrow \lim\limits_{n\to\infty} P_n((\varphi \wedge \psi) \vee (\varphi \wedge \neg\psi) \to \varphi \wedge \neg\psi) = \lim\limits_{n\to\infty} P_n(\varphi \to \varphi \wedge \neg\psi) = \lim\limits_{n\to\infty} P_n(\varphi \to \neg\psi) = 0$。(因此$P_n(\varphi)$只有有限个值可能为0。若它有无穷个值为0，则$P_n(\varphi \to \neg\psi)$中有无穷个值为1，从而该极限不能为0。因此不失一般性地假设$P_n(\varphi)$的所有值为正，此时，

$P_n(\varphi \to \neg\psi) = 1 - P(\varphi \to \psi) \Leftrightarrow \lim\limits_{n\to\infty} P_n(\varphi \to \psi) = 1$.

情况2．据定义5.9，$\varphi \leqslant \bot \Leftrightarrow \lim\limits_{n\to\infty} P_n(\varphi \vee \bot \to \bot) > 0$，因为$P_n(\varphi \vee \bot \to \bot)$只可能为0或1：当$P_n(\varphi) = 0$时，$P_n(\varphi \vee \bot \to \bot) = 1$；当$P_n(\varphi) > 0$时，$P_n(\varphi \vee \bot \to \bot) = 0$，所以$P_n(\varphi)$的值中有无穷个值等于0，因此

$$\lim_{n\to\infty} P_n(\varphi \to \psi) = 1.$$

反之，若$\lim\limits_{n\to\infty} P_n(\varphi \to \psi) = 1$，这里又分两种情况讨论。情况(i)：若$P_n(\varphi)$有无穷个值不为0，不失一般性假设$P_n(\varphi)$的值皆为正，因此从$\lim\limits_{n\to\infty} P_n(\varphi \to \psi) = 1$可得

$$\lim_{n\to\infty} P_n(\varphi \to \neg\psi) = 0,$$

据上面情况1的前半部可得$\varphi \wedge \neg\psi < \varphi \wedge \psi$，据定义5.8(2)，$\varphi \to \psi$在≤中成立。情况(ii)：若$P_n(\varphi)$有无穷个值

为 0，不失一般性，设 $P_n(\varphi)$ 的所有值皆为 0，则对所有 P_n，$P_n(\varphi\vee\bot)=0$，所以 $\lim\limits_{n\to\infty}P_n(\varphi\vee\bot\to\bot)=1>0$. 据定义 5.9，$\varphi\leqslant\bot$，所以据定义 5.8(2)，$\varphi\to\phi$在$\leqslant$中成立.

证(4) 借助定理 5.6(2)，(3)来证. 假设存在一 P 偏序$\leqslant$使得所有 $B\in S$ 在$\leqslant$中成立而 A 在$\leqslant$中不成立，则据定理 5.6(2)，存在与$\leqslant$关联的齐一概率函数序列 P_1，P_2，$\cdots$. 由定理 5.6(3)，对所有 $B\in S$，$\lim\limits_{n\to\infty}P_n(B)=1$，但 $\lim\limits_{n\to\infty}P_n(A)<1$，因此 $S\nvdash A$，所以据定理 5.2(2)，$S\not\Vdash A$.

§4 公式集的标准偏序

前面我们已经建立了下列定理：(1) 若 $S\vdash_p A$，则 $S\Vdash A$. (2) 若 $S\Vdash A$，则 A 在所有使 S 的全部公式成立的 P 偏序中成立.

现在我们来建立一种完全性定理：若 A 在所有使 S 的全部公式成立的 P 偏序中成立，则 $S\vdash_p A$. 这个定理将在下一节证明，但这个证明依赖一个预备性定理，即关于有穷公式集 S（若 S 一致）的标准 P 偏序定理. 这对给出确定一公式 A 是否是一有穷公式集 S 的合理后承的判定程序也是有用的.

定义 5.10 令 $\mathscr{L}$ 有穷，$S\subseteq_f\mathscr{L}$，我们有

(1) S 的直接归约 $\mathrm{Red}(S)=\{A\in S:\mathrm{SD}(A)\subseteq\boldsymbol{SD}(S)\}$.

(2) S 的归约序列是 S 的一子集序列 $S_1,\cdots,S_p$ 使得对 $i=1,\cdots,p$，$S_1=S$，$S_{i+1}=\mathrm{Red}(S_i)$，对 $i=1,\cdots,p-1$，S_{i+1}是S_i 的真子集，且 $\mathrm{Red}(S_p)=S_p$.

(3) 由 S 生成的 $\mathrm{SD}\mathscr{L}$ 的有序划分是 $\mathrm{SD}\mathscr{L}$ 的子集序列 $\mathrm{SD}\mathscr{L}_1,\cdots,\mathrm{SD}\mathscr{L}_{p+1}$ 且

$$\mathrm{SD}\mathscr{L}_1=\mathrm{SD}\mathscr{L}-\boldsymbol{SD}(S_1),$$

$$\mathrm{SD}\mathscr{L}_2 = \boldsymbol{SD}(S_1) - \boldsymbol{SD}(S_2), \cdots,$$
$$\mathrm{SD}\mathscr{L}_{i+1} = \boldsymbol{SD}(S_i) - \boldsymbol{SD}(S_{i+1}),$$
$$\mathrm{SD}\mathscr{L}_{p+1} = \boldsymbol{SD}(S_p),$$
其中 $i = 1, \cdots, p-1, S_1, \cdots, S_p$ 是 S 的归约序列.

(4) 若 $\boldsymbol{SD}(S) \neq \mathrm{SD}\mathscr{L}$ 且由 S 生成的 $\mathrm{SD}\mathscr{L}$ 的有序划分是 $\mathrm{SD}\mathscr{L}_1, \cdots, \mathrm{SD}\mathscr{L}_{p+1}$,则与 S 关联的 $\mathscr{L}$ 的标准 P 偏序是 P 偏序$\leqslant$使得对所有α, $\beta \in \mathrm{SD}\mathscr{L}$, 若 $\alpha \in \mathrm{SD}\mathscr{L}_i$, $\beta \in \mathrm{SD}\mathscr{L}_j$, 则 $\alpha \leqslant \bot \Leftrightarrow i = p+1, \alpha \leqslant \beta \Leftrightarrow j \leqslant i$.

注意(4)假设了未证明的命题: 由 S 生成的 $\mathrm{SD}\mathscr{L}$ 的有序划分是 $\mathrm{SD}\mathscr{L}$ 的划分. 有关划分和与之关联的 P 偏序的重要性质,我们将在后面推出. 现在我们先看一个简单的例子,它有助于我们在直观上把握这些概念.

语言 $\mathscr{L}$ 只由两原子句"p"和"q"(再加上"$\top$"和"$\bot$")生成. $\mathrm{SD}\mathscr{L}$ 可以看作由下列四个公式组成: $a=$"$\neg p \wedge \neg q$", $b=$"$p \wedge \neg q$", $c=$"$p \wedge q$", $d=$"$\neg p \wedge q$". 令 $S = \{p \to q, \neg q \to \neg p, q \to \neg p, p \vee q \to p\}$. 为了确定 $\mathrm{Red}(S)$, 我们必须确定 S 中哪个公式 A 有性质 $\mathrm{SD}(A) \subseteq \boldsymbol{SD}(S)$. 从定义 5.3 得知

$$\mathrm{SD}(A) = \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \mathrm{Ant}(A)\},$$
$$\boldsymbol{SD}(A) = \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \mathrm{Ant}(A) \wedge \neg \mathrm{Cons}(A)\},$$
$$\boldsymbol{SD}(S) = \bigcup_{A \in S} \boldsymbol{SD}(A).$$

如下表所示.

	$A \in S$	$\mathrm{SD}(A)$	$\boldsymbol{SD}(A)$
1.	$p \to q$	$\{b, c\}$	$\{b\}$
2.	$\neg q \to \neg p$	$\{a, b\}$	$\{b\}$
3.	$q \to \neg p$	$\{c, d\}$	$\{c\}$
4.	$p \vee q \to p$	$\{b, c, d\}$	$\{d\}$

$$SD(S)=\{b,c,d\}.$$

显然，使得 $\mathrm{SD}(A)\subseteq SD(S)$ 的 S 的公式 A 是 $p\to q, q\to\neg p$, $p\vee q\to p$，因此

$$\mathrm{Red}(S)=\{p\to q, q\to\neg p, p\vee q\to p\}.$$

为了构造 S 的归约序列，我们简单重述上述过程。令 $S_1=S$，$S_2=\mathrm{Red}(S)=\{p\to q, q\to\neg p,\ p\vee q\to p\}$，$S_3=\mathrm{Red}(S_2)=S_2$，因此这个归约序列终止于 S_2。因此由 S 生成的 $\mathrm{SD}\mathscr{L}$ 的有序划分如下构造：

$\mathrm{SD}\mathscr{L}_1=\mathrm{SD}\mathscr{L}-SD(S_1)=\{a,b,c,d\}-\{b,c,d\}=\{a\}$,
$\mathrm{SD}\mathscr{L}_2=SD(S_1)-SD(S_2)=\{b,c,d\}-\{b,c,d\}=\phi$,
$\mathrm{SD}\mathscr{L}_3=SD(S_2)=\{b,c,d\}$.

给定这样的有序划分，与 S 关联的 $\mathscr{L}$ 的标准 P 偏序的构造本质上就是倒过来排列这个划分中的元素，且使 $\mathrm{SD}\mathscr{L}_3$ 的所有元素等价 $\bot$，即这个标准偏序由下列 $\mathrm{SD}\mathscr{L}\cup\{\bot\}$ 上的偏序确定：

$$\bot\approx b\approx c\approx d<a.$$

对所有 $\varphi,\phi\in\mathscr{L}^*$，$\mathscr{L}^*$ 上的偏序关系由上述状态描述的偏序确定，这一点在定理5.5已说清楚：φ 和 ϕ 的偏序关系等同于 $\mathrm{SD}\mathscr{L}^+$ 中分别极大蕴涵 φ 和 ϕ 的状态描述的偏序关系。因此为了确定 p 和 q 的偏序关系，我们只须注意所有蕴涵它们的状态描述等价，从而 p 和 q 也等价。另一方面，极大蕴涵 p 的状态描述严格小于极大蕴涵 $\neg p$ 的状态描述，因此 p 在这个偏序中也严格小于 $\neg p$。特别是，在该标准偏序中有，$p\leqslant\bot$，$\neg q\wedge\neg(\neg p)<\neg q\wedge\neg p$，$q\leqslant\bot$，$p\vee q\leqslant\bot$.

由此可得 S 的四个公式在该标准偏序中（在定义5.8(2)的意义上）全都成立。这个事实决非偶然，关于与一公式集 S 关联的标准偏序的一个最基本的事实，就是 S 的全部公式在该偏序中成立。我们在下列定理中证明这一点。

定理5.7 令$\mathscr{L}$有穷，$S \subsetneq_f \mathscr{L}$，$SD\mathscr{L}_1, \cdots, SD\mathscr{L}_{p+1}$是由$S$生成的$SD\mathscr{L}$的有序划分，则

(1) $SD\mathscr{L}_1, \cdots, SD\mathscr{L}_{p+1}$是$SD\mathscr{L}$的一划分.

(2) 若$\boldsymbol{SD}(S) \neq SD\mathscr{L}$，则$S$的全部公式在与$S$关联的标准$P$偏序中成立.

证明：(1) 该定理直接从以下事实得到：在S的归约序列$S_1, \cdots, S_p$中的每一元素是前一元素的子集，即$S_p \subseteq \cdots \subseteq S_1$，因此子集关系$\boldsymbol{SD}(S_p) \subseteq \boldsymbol{SD}(S_{p-1}) \subseteq \cdots \subseteq \boldsymbol{SD}(S_1)$成立. 从初等集合论可得$SD\mathscr{L}_1 = SD\mathscr{L} - \boldsymbol{SD}(S_1), \cdots, SD\mathscr{L}_{p+1} = \boldsymbol{SD}(S_p)$互不相容，且它们的并就是$SD\mathscr{L}$.

(2) 对所有$A \in S$，假设$SD\mathscr{L} - \boldsymbol{SD}(S) \neq \varnothing$，因此与$S$关联的标准$P$偏序$\leqslant$被定义. 不失一般性地假设$A$是条件句，因为若$A \in \mathscr{L}^+$，则$A$在$\leqslant$中成立$\Longleftrightarrow$条件句公式$\mathrm{Ant}(A) \to \mathrm{Cons}(A) = \mathrm{Cond}(A)$在$\leqslant$中成立(据定义5.2(2)，$\mathrm{Ant}(A) = \top$). 这种方法我们在以后的定理证明中将经常用到.

假设$A = \varphi \to \psi$：我们要证A在$\leqslant$中成立. 首先据定义5.10(3)，从由S生成的有序划分$SD\mathscr{L}_1, \cdots, SD\mathscr{L}_{p-1}$直接可得

$\boldsymbol{SD}(S_i) = SD\mathscr{L}_{i-1} \cup \cdots \cup SD\mathscr{L}_{p+1}$，$i = 1, 2, \cdots, p$，其中$S_1, \cdots, S_i, \cdots, S_p$是$S$的归约序列. 因为$A \in S = S_1$，所以$A$至少在该归约序列的某个$S_i$中. 先假设$A \in S_p$，则

$$SD(A) \subseteq \boldsymbol{SD}(S_p) = SD_{p+1},$$

而据定义5.10(4)，$SD\mathscr{L}_{p+1} = \{\alpha \in SD\mathscr{L} : \alpha \text{在} \leqslant \text{中等价} \perp\}$，据(5.4)，$\varphi \leqslant \perp$.因此据定义5.8(2)，$\varphi \to \psi$在$\leqslant$中成立. 假设$A \notin S_p$，则存在某个$i$使得$A \in S_i$但$A \notin S_{i+1}$.

因为$A \in S_i$，所以据定义5.3(3)，$\boldsymbol{SD}(A) \subseteq \boldsymbol{SD}(S_i)$，因

此 $\boldsymbol{SD}(A) \subseteq \mathrm{SD}\mathscr{L}_{i+1} \cup \cdots \cup \mathrm{SD}\mathscr{L}_{p+1}$，亦即 $\boldsymbol{SD}(A) = \{\alpha \in \mathrm{SD}\mathscr{L}_{i+1} \cup \cdots \cup \mathrm{SD}\mathscr{L}_{p+1} : \alpha \models \varphi \wedge \neg\psi\}$. 从而在 $\mathrm{SD}\mathscr{L}_{i+1} \cup \cdots \cup \mathrm{SD}\mathscr{L}_{p+1}$ 中必存在一元素 β 极大蕴涵 $\varphi \wedge \neg\psi$.

因为 $A \notin S_{i+1} - \mathrm{Red}(S_i)$，所以据定义 5.10(1)，$\mathrm{SD}(A) \nsubseteq \boldsymbol{SD}(S_i)$，因此 $\mathrm{SD}(A)$中必包含一 $\alpha \in \mathrm{SD}\mathscr{L}_1 \cup \cdots \cup \mathrm{SD}\mathscr{L}_i$，且使得 α 极大蕴涵 $\varphi \wedge \psi$（因为 $\alpha \in \mathrm{SD}\mathscr{L}_1 \cup \cdots \cup \mathrm{SD}\mathscr{L}_i$ 且 $\boldsymbol{SD}(A) \subseteq \mathrm{SD}\mathscr{L}_{i+1} \cup \cdots \cup \mathrm{SD}\mathscr{L}_{p+1}$，所以 $\alpha \models \varphi$ 且 $\alpha \models \psi$. 若 $\alpha' \models \varphi \wedge \psi$，则 $\alpha' \leqslant \alpha$).

因此据定义 5.10(4)，$\beta < \alpha$. 由(5.4)，$\varphi \wedge \neg\psi < \varphi \wedge \psi$. 从而据定义 5.8(2)，$\varphi \rightarrow \psi$ 在$\leqslant$中成立.

§5 完全性定理和判定程序

定理 5.8 令 $\mathscr{L}$ 有穷，$A \in \mathscr{L}$，$S \underset{f}{\subseteq} \mathscr{L}$，$S' = S \cup \{\sim A\}$，$S'_1, \cdots, S'_p$ 是 S' 的归约序列，$\mathrm{SD}\mathscr{L}'_1, \cdots, \mathrm{SD}\mathscr{L}'_{p+1}$ 是 S' 生成的 $\mathrm{SD}\mathscr{L}$ 的有序划分，$S_0 = S'_p - \{\sim A\}$，则下列条件等价.

(1) $S \underset{p}{\vdash} A$.

(2) $S \Vdash A$.

(3) A 在所有使 S 的全部公式成立的 P 偏序中成立.

(4) $\mathrm{SD}(A) \subseteq \mathrm{SD}(S'_p)$.

(5) $\boldsymbol{SD}(A) \subseteq \boldsymbol{SD}(S_0)$，且 $\mathrm{SD}(S_0) - \boldsymbol{SD}(S_0) \subseteq \mathrm{SD}(A) - \boldsymbol{SD}(A)$.

(6) 对某个 $S'' \subseteq S$，$\boldsymbol{SD}(A) \subseteq \boldsymbol{SD}(S'')$，且 $\mathrm{SD}(S'') - \boldsymbol{SD}(S'') \subseteq \mathrm{SD}(A) - \boldsymbol{SD}(A)$.

证明：(1)⇒(2) 和 (2)⇒(3) 在定理 5.3 和定理 5.6 已证. (5)⇒(6) 显然. 剩下要证的是 (3)⇒(4)⇒(5) 和 (6)⇒

(1).

(3)⇒(4). 设(3)成立. 我们在此总假设A是条件句$\varphi\to\psi$,因为若A是某个非条件句φ,则它可以被$\top\to\varphi$置换,并不改变我们将要考虑的任何条件. 现考虑

$$S'=S\cup\{\sim A\},$$

其中$\sim A$如定义5.2(1)定义: $\sim A=\varphi\to\neg\psi$.

情况1. $\boldsymbol{SD}(S')=\mathrm{SD}\mathscr{L}$, 在这种情况下与$S'$关联的标准$P$偏序无定义,显然条件(4)成立,因为

$$\mathrm{SD}(A)\subseteq\mathrm{SD}\mathscr{L}=\boldsymbol{SD}(S'),$$

以及对所有$B\in S'$,$\mathrm{SD}(B)\subseteq\mathrm{SD}\mathscr{L}=\boldsymbol{SD}(S')$, 所以

$$S_2'=\mathrm{Red}(S_1')=\mathrm{Red}(S')$$
$$=\{B\in S':\mathrm{SD}(B)\subseteq\boldsymbol{SD}(S')=\mathrm{SD}\mathscr{L}\}=S_1'.$$

因此这个序列终至于S_1'(即$p=1$). 所以

$$\mathrm{SD}\mathscr{L}_1=\boldsymbol{SD}(S_1')=\mathrm{SD}\mathscr{L},$$

亦即

$$\mathrm{SD}(A)\subseteq\boldsymbol{SD}(S')=\boldsymbol{SD}(S_1')=\boldsymbol{SD}(S_p')\subseteq\mathrm{SD}(S_p').$$

情况2. $\boldsymbol{SD}(S')\neq\mathrm{SD}\mathscr{L}$, 据定义5.10(4), 能够确定与$S'$关联的标准$P$偏序$\leqslant$. 因为$\sim A\in S'$,且据定理5.7(2), S'的所有公式在$\leqslant$中成立,且$S\subseteq S'$,所以S的所有公式在$\leqslant$中成立. 据条件(3), A也在$\leqslant$中成立. 从定义5.8(2), 仅当$\mathrm{Ant}(A)=\mathrm{Ant}(\sim A)=\varphi\leqslant\bot$时, $A=\varphi\to\psi$和$\sim A=\varphi\to\neg\psi$才能同在一$P$偏序$\leqslant$中成立.

因为A在$\leqslant$中成立$\Longleftrightarrow\varphi\wedge\neg\psi<\varphi\wedge\psi$或$\varphi\leqslant\bot$,$\sim A$在$\leqslant$中成立$\Longleftrightarrow\varphi\wedge\psi<\varphi\wedge\neg\psi$或$\varphi\leqslant\bot$,所以仅当$\varphi\leqslant\bot$时,两者皆成立. 条件(4)可以从$\varphi\leqslant\bot$直接得到. 据定义5.10(4)和$S_p'$的构造,若$\varphi\leqslant\bot$能在该标准序列中成立,则$\{\alpha\in\mathrm{SD}\mathscr{L}:\alpha\models\varphi\}\subseteq\mathrm{SD}\mathscr{L}'_{p+1}$(证明: 设任意$\alpha\in\{\alpha\in\mathrm{SD}\mathscr{L}:\alpha\models\varphi\}$, 则$\alpha\models\varphi$, 据定理5.5的(5.4), $\varphi\leqslant\bot\Longleftrightarrow\alpha'\leqslant\bot$,

其中α'是φ的极大蕴涵元．因为$\alpha\leqslant\alpha'$，所以$\alpha\leqslant\bot$，据定义 5.10(4)，$\alpha\in \mathrm{SD}\mathscr{L}_{p+1}$)．而据定义 5.10(3)，

$$\mathrm{SD}\mathscr{L}'_{p+1}=\boldsymbol{SD}(S'_p),\mathrm{SD}(A)=\{\alpha\in\mathrm{SD}\mathscr{L}:\alpha\models\varphi\},$$

所以 $\mathrm{SD}(A)\subseteq\boldsymbol{SD}(S'_p)$．因为 $\boldsymbol{SD}(S'_p)\subseteq\mathrm{SD}(S'_p)$，所以

$$\mathrm{SD}(A)\subseteq\mathrm{SD}(S'_p).$$

(4)⇒(5) 假设(4)成立．我们分别考虑两种情况．

情况 1．$\sim A\notin S'_p$，因为(4)成立，所以

$$\mathrm{SD}(A)\subseteq\mathrm{SD}(S'_p).$$

又因为 $\boldsymbol{SD}(A)\subseteq\mathrm{SD}(A)$，所以 $\boldsymbol{SD}(A)\subseteq\mathrm{SD}(S'_p)$．据定义 5.10(2)，$S'$的归约序列是 $S'_1,\cdots,S'_p$，且

$$\mathrm{Red}(S'_p)=S'_p=\{B\in S':\mathrm{SD}(B)\subseteq\boldsymbol{SD}(S'_p)\},$$

因此 $\mathrm{SD}(S'_p)\subseteq\boldsymbol{SD}(S'_p)$，所以 $\mathrm{SD}(S'_p)=\boldsymbol{SD}(S'_p)$，所以 $\boldsymbol{SD}(A)\subseteq\boldsymbol{SD}(S'_p)=\boldsymbol{SD}(S_0)$．因为 $S_0=S'_p$，$\mathrm{SD}(S'_p)=\boldsymbol{SD}(S'_p)$，所以 $\mathrm{SD}(S_0)-\boldsymbol{SD}(S_0)=\varnothing$，所以 $\mathrm{SD}(S_0)-\boldsymbol{SD}(S_0)\subseteq\mathrm{SD}(A)-\boldsymbol{SD}(A)$．

情况 2．$\sim A=\varphi\to\neg\psi\in S'_p$．假设 $\alpha\in\boldsymbol{SD}(A)$，即 $\alpha\models\varphi\wedge\neg\psi$．因为 $\boldsymbol{SD}(A)\subseteq\mathrm{SD}(A)$，所以$\alpha\in\mathrm{SD}(A)$，据条件(4)，$\alpha\in\mathrm{SD}(S'_p)=\boldsymbol{SD}(S'_p)$．因为 $S'_p=S_0\cup\{\varphi\to\neg\psi\}$，所以 $\boldsymbol{SD}(S'_p)=\boldsymbol{SD}(S_0)\cup\boldsymbol{SD}(\varphi\to\neg\psi)$，所以 $\alpha\in\boldsymbol{SD}(S_0)\cup\boldsymbol{SD}(\varphi\to\neg\psi)$，但 $\boldsymbol{SD}(\varphi\to\neg\psi)=\{\alpha\in\mathrm{SD}\mathscr{L}:\alpha\models\varphi\wedge\psi\}$，所以 $\alpha\notin\boldsymbol{SD}(\varphi\to\neg\psi)$，因此 $\alpha\in\boldsymbol{SD}(S_0)$，所以 $\boldsymbol{SD}(A)\subseteq\boldsymbol{SD}(S_0)$．下面证

$$\mathrm{SD}(S_0)-\boldsymbol{SD}(S_0)\subseteq\mathrm{SD}(A)-\boldsymbol{SD}(A).$$

假设 $\alpha\in\mathrm{SD}(S_0)-\boldsymbol{SD}(S_0)$，因为 $S_0\subseteq S'_p$，所以

$$\alpha\in\mathrm{SD}(S'_p)=\boldsymbol{SD}(S'_p)=\boldsymbol{SD}(S_0)\cup\boldsymbol{SD}(\varphi\to\neg\psi),$$

但$\alpha\notin\boldsymbol{SD}(S_0)$，所以

$$\alpha\in\boldsymbol{SD}(\varphi\to\neg\psi)=\boldsymbol{SD}(\sim A)$$
$$=\{\alpha'\in\mathrm{SD}\mathscr{L}:\alpha'\models\varphi\wedge\psi\}$$

$= \mathrm{SD}(A) - \boldsymbol{SD}(A)$.

(6)⇒(1). 假设 (6) 成立,即对某个 $S'' \subseteq S$, $\boldsymbol{SD}(A) \subseteq \boldsymbol{SD}(S'')$ 且 $\mathrm{SD}(S'') - \boldsymbol{SD}(S'') \subseteq \mathrm{SD}(A) - \boldsymbol{SD}(A)$. 假设 $S'' = \{\varphi_1 \to \psi_1, \cdots, \varphi_n \to \psi_n\}$. 考虑下列公式

$$B = (\varphi_1 \vee \cdots \vee \varphi_n) \to (\neg(\varphi_1 \wedge \neg\psi_1) \wedge \cdots \wedge \neg(\varphi_n \wedge \neg\psi_n)).$$

据定理 5.3(4), 因为 B 的前件是 $\varphi_1 \to \psi_1, \cdots, \varphi_n \to \psi_n \in S''$ 的前件的析取,所以

$$\mathrm{SD}(B) = \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \varphi_1 \vee \cdots \vee \varphi_n\}$$

$$= \bigcup_{\varphi_i \to \psi_i \in S''} \mathrm{SD}(\varphi_i \to \psi_i) = \mathrm{SD}(S''),$$

所以 $\mathrm{SD}(B) = \mathrm{SD}(S'')$, 因为 $\mathrm{Ant}(B) \wedge \neg\mathrm{Cons}(B)$ 重言等价 $(\varphi_1 \wedge \neg\psi_1) \vee \cdots \vee (\varphi_n \wedge \neg\psi_n)$, 所以

$$\boldsymbol{SD}(B) = \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \mathrm{Ant}(B) \wedge \neg\mathrm{Cons}(B)\}$$

$$= \bigcup_{\varphi_i \to \psi_i \in S''} \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \varphi_i \wedge \neg\psi_i\} = \boldsymbol{SD}(S''),$$

所以 $\boldsymbol{SD}(B) = \boldsymbol{SD}(S'')$, 因此

$$\mathrm{SD}(B) - \boldsymbol{SD}(B) = \mathrm{SD}(S'') - \boldsymbol{SD}(S'').$$

下列命题易证: 对任意 $A, B \in \mathscr{L}$,

$$\{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models A\} \subseteq \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models B\} \Longleftrightarrow A \models B. \quad (5.7)$$

现在据条件 (6),

$$\boldsymbol{SD}(A) = \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \mathrm{Ant}(A) \wedge \neg\mathrm{Cons}(A)\}$$
$$\subseteq \boldsymbol{SD}(S'') = \boldsymbol{SD}(B) = \{\alpha \in \mathrm{SD}\mathscr{L} : \alpha \models \mathrm{Ant}(B) \wedge \neg\mathrm{Cons}(B)\},$$

因此据命题 (5.7),

$$\mathrm{Ant}(A) \wedge \neg\mathrm{Cons}(A) \models \mathrm{Ant}(B) \wedge \neg\mathrm{Cons}(B),$$

又据 (6)

$$\mathrm{SD}(B) - \boldsymbol{SD}(B) = \mathrm{SD}(S'')$$

$$-\boldsymbol{SD}(S'')\subseteq \mathrm{SD}(A)-\boldsymbol{SD}(A),$$

即

$$\boldsymbol{SD}(\sim B)\subseteq \boldsymbol{SD}(\sim A),$$

因此

$$\{\alpha\in \mathrm{SD}\mathscr{L}:\alpha\models \mathrm{Ant}(B)\wedge \mathrm{Cons}(B)\}$$

$$\subseteq\{\alpha\in \mathrm{SD}\mathscr{L}:\alpha\models \mathrm{Ant}(A)\wedge \mathrm{Cons}(A)\},$$

据命题(5.7),

$$\mathrm{Ant}(B)\wedge \mathrm{Cons}(B)\models \mathrm{Ant}(A)\wedge \mathrm{Cons}(A),$$

据定理 5.3(5)，从

$$\mathrm{Ant}(A)\wedge \neg\mathrm{Cons}(A)\models \mathrm{Ant}(B)\wedge \neg\mathrm{Cons}(B)$$

和

$$\mathrm{Ant}(B)\wedge \mathrm{Cons}(B)\models \mathrm{Ant}(A)\wedge \mathrm{Cons}(A)$$

可得 $B\mid\!\!\underset{p}{-}A$，而据定理 5.3(4)，$S\mid\!\!\underset{p}{-}B$，所以 A 属于 S 的概率后承集，即 $S\mid\!\!\underset{p}{-}A$。证毕.

注意：定理 5.8 提供了一个判定程序，来确定一公式 A 是否是一有穷公式集 S 的概率后承. 这种判定程序就是构造 $S'=S\cup\{\sim A\}$ 的归约序列，并确定 SD(A) 是否是 SD(S'_p) 的子集 (其中 S'_p 是 S' 的归约序列中最后一项)，据定理 5.8,

$$S\mid\!\!\underset{p}{-}A\Longleftrightarrow \mathrm{SD}(A)\subseteq \mathrm{SD}(S'_p).$$

根据完全性定理，我们能得到一些有趣且直观意义清楚的结果，请参见 [12].

Adams 的系统尽管在逻辑上已经完备化(即有了完全性定理)，许多概念的引入和定理的证明技巧也独具魅力，但从整体上看仍然非常有限，这表现在许多主要结果中. 例如在完全性定理中不仅要求初始命题(原子命题)集有穷，还要求前提集 (S) 有穷. 这些条件从逻辑的角度看当然是非常苛刻的. 这也说明用有穷逻辑(合取，析取和量化的范围有穷)刻划概率概念是非常受限制的. 这样的困难直到最后一章才被无穷概率逻辑克服.

第六章　模态归纳逻辑(上)

——Cohen的归纳支持逻辑

L. J. Cohen 的归纳逻辑是一种模态归纳逻辑. 我们知道,模态逻辑可以说是研究可能的逻辑,归纳逻辑和概率逻辑则进一步研究这种可能性的程度. 从这个意义上说, 模态逻辑和归纳逻辑、概率逻辑是有一定的联系的(历史上就有人用概率语义学来刻划模态逻辑). Cohen 和 Burks 在这个方面做了有益的探索.

Cohen 的理论既是对 F. Bacon 和 J. S. Mill 的消除归纳法的概括, 又是对 Keynes 关于归纳概率是论证的可靠性的量度的思想的发展. 另一方面, Cohen 从科学哲学的高度,对实验科学中的科学方法进行抽象, 提出了相干变量法, 并且力图用此来解释模态逻辑 S_4 的一概括的公理化形式系统,因此 Cohen 的工作开辟了研究归纳逻辑的新领域.

§1　对概率概念的哲学思考

1. 概率理论是多元标准的

Cohen 认为, 概率的哲学理论是揭示产生各种概然命题表面多样性背后的统一性. 在日常生活中我们常遇到明显不同的概率类型. 例如, 从一付洗好的牌中抽取一张花点的牌的概率、货车司机活到 70 岁的概率、一镭原子在 24 小时内将要衰变的概率、一体育明星在未来比赛中取胜的概率等等.过

去概率的哲学理论企图把所有这样的概率解释为一种单一标准的概率。例如，Reichenbach 等人的相对频率、L. J. Savage 等人的信念度、Carnap 等人的命题之间的逻辑关系等等。

但是这样的概率理论难以解释以上各种概然命题。因此近几十年，一直存在一种基本倾向，即赞成概率是多元标准的观点。这种观点认为，这样或那样的单一标准，若它们自身一致，都不会让多样化的事实完全驳倒，只是限制在它们各自的运用范围中。K. Popper 对此做出了重要贡献，他证明了有可能构造一个概率的数学演算来作为纯形式系统，除了与该系统的公理一致的逻辑语法外，对概率函数的本性不作任何别的假设。Cohen 把这种概率称为数学概率，他认为这样的理论虽然可以容纳多元标准，但本身也遇到不少麻烦，主要是不能澄清为何对概率概念有多种多样的语义标准，而且每一个都以自己的标准满足概率演算的公理以及都有各自的实际效用。

2. 概率就是对可证性的分级（gradation）

Cohen自己提出另一种概率的多元标准，这种标准可以避免纯语法的形式主义，具有数学概率概念不具备的解释能力。

假设H是一全称概括条件句，对H的可信赖性（reliability）就有一个一元的归纳支持的等级 $S(H)$（这个概念后面还要详细定义和分析），因此从H的前件到后件就相应有一证明。一元概率 $P(H)$ 就是从H的前件到后件的可证性等级，也就是从前件到后件的推理可靠性（soundness）的程度，所以 Cohen 认为，$S(H)=P(H)$。假定H是$(x_1)(x_2)\cdots(x_n)(R\rightarrow S)$，其中 $x_1, x_2, \cdots, x_n$ 是R和S中所有自由出现的个体变元，则相应的二元函数应是 $P(S(a_1,a_2,\cdots,a_n),R(a_1,a_2,\cdots,a_n))=S((x_1)(x_2)\cdots(x_n)(R\rightarrow S))$，其中 $a_1,a_2,\cdots,a_n$ 是某个适当论域中的个体常元，$R(a_1, a_2, \cdots,$

$a_n) \rightarrow S(a_1, a_2, \cdots, a_n)$是$H$的代入特例。

例如，我们根据消除归纳法，在实验证据中，对概括句“所有的蜜蜂有色感”建立一适当的等级，从而对命题“若这个昆虫是蜜蜂，则它有色感”也建立一适当的等级，那么我们也已经从前提“这个昆虫是蜜蜂”对结论“这个昆虫有色感”的可证性建立了同样的等级。因此 Cohen 把概率看作是对推理而不是陈述句的评价(而归纳支持则相反)。

Cohen 认为他的把概率看作是可证性分级的观点是多元标准的。而前面提到的那些数学概率可以看作对应不同类型的可证性。例如概率的相对频率概念可以看作类似于一个推演规则是一般、外延和偶然的（经验的）演绎系统中的可证性，而概率的逻辑关系的概念却类似于一个推演规则是单称、必然的演绎系统中的可证性。既然不同的证明标准对应不同的可证性的陈述句，概率的不同概念之间的主要区别就可以由上述事实得到解释：可证性也有不同的种类。这样各种单一概率标准之间的主要差异就通过对可证性的种类的考察被揭示出来。

3. 数学概率与归纳概率

先引入 Cohen 理论的几个重要的术语。$P(\ )$和$P(\cdots, -)$分别称为一元概率函子（functor）和二元概率函子。Cohen 把函子看作是一指谓函数的表达式，函子中变目位置上可以填入适当的语句或谓式(predicable)，它区别于我们通常使用的 predicate，前者是一表达式(注意，Cohen 的表达式不是句子而是词组)，它起的作用是作为一句子的逻辑谓词（“predicate”是表达式中起谓述作用的一实际出现，所以为了区别起见，姑且把前者译为“谓式”)。因为他的概率是多元标准的，所以变目位置可以根据可证性的种类填入句子(象 Carnap 逻辑)或谓式(象 Reichenbach 逻辑)。

Cohen 把概率概念分成数学概率概念 (P_M) 和归纳概率概念 (P_I)，数学概率概念在结构上和机会的数学演算 (the mathematical calculus of chance)的原则是一致的。例如合取的乘法原则：$P(A \wedge B, C) = P(A, C) \cdot P(B, A \wedge C)$ 和否定的补余原则：$P(\neg B, A) = 1 - P(B, A)$。Cohen 认为完全性作为某些演绎系统的性质，可以看作是概率补余性的极限情况。一个系统称为完全的，如果任一公式 B 从公理可证 $\Leftrightarrow \neg B$ 不可证.若概率表达为推理可靠性的程度,则从 A 证明 B 就是概率的极限情况,即 $P(B, A) = 1$。若把这种极限情况看作是可证性在完全系统的特例,则我们可以用$P(\neg B, A) = 0$来描述"从 $A, \neg B$ 不可证"的情况。

归纳概率的概念和数学概率的概念相比，在结构上有很大的不同.例如它有完全不同的否定原则和合取原则(这个我们后面还要仔细论述),因为它涉及的是概率的可比较的分级或有序的分级,而不是定量的或可测度的分级,因此它特别适用于证据不可计算或不可测度的推理领域.

Cohen 认为把概率看作是可证性的程度，不仅能说明哲学上熟悉的各种各样的概率形式(Reichenbach, Carnap, Savage 等)，而且说明了它们通常与数学上的概率演算的原则是一致的。Cohen 认为他的概率概念还可以用于发现迄今未加分析的概率概念，并且能说明这些概念能满足一些相当不同的语法原则,特别是使 $P(\neg B, A) = 0$ 和 $P(B, A) = 0$ 并存,并且当 $P(B, A) > 0$ 时,通常有 $P(\neg B, A) = 0$。也就是说，在 Cohen 可证性的意义上,还有许多证明系统是不完全的,使得在类似的概率系统中,有可能同时有 $P(A, B) = 0$ 且 $P(\neg A, B) = 0$。

这里需要说明的是,尽管Cohen 指出了数学概率的不足，尤其是它不能用于科学理论证明（证据对全称命题的假说的

支持)和英美法律体系的司法推理程序等领域,但他并没有抹杀数学概率的作用，认为它在自己适用的领域有着别的概率概念不可取代的作用. 在[13],他提到：本书的中心论题是，归纳概率概念作为人类理性一个长期有效的工具，是和数学概率概念并存的. 它们中任何一个都不能根据任何可能的逻辑分析方式归结为另一个. 在我们实际的推理中，那一个也不能永久地取代另一个,因为它们担负着不同的任务.

§2 Cohen 的归纳逻辑的语义理论

1. 归纳支持与归纳概率

为了给归纳概率的概念提供合理重建，就有必要首先对归纳支持概念提供合理重建. 在科学实践中，人们往往提出一个或多个全称概括句形式的假说，然后通过观察和实验搜集支持假说的证据. 证据对假说的支持往往不是演绎关系，我们称这种支持为归纳支持.

归纳概率不是归纳支持的特殊形式，区别这一点是很重要的. 粗略说，一个一阶单称命题 S 根据另一证据 R 的归纳概率是和一可检验的全称条件量化句 $(x)(R(x)\to S(x))$ 的归纳支持的等级相同的. 因此，归纳支持本质上关注全称条件量化句，而归纳概率则关注这样的概括句（相对它们的前件）的代入特例的后件. 归纳概率的陈述句的逻辑语法在几个重要方面不同于归纳支持的陈述句的逻辑语法.

考虑可以从关于事件的全称条件句得到的那些推理规则. 例如,规则

(1) 从乌云出现推断即将下雨.

可以从以下全称条件句得到:

(2) 对于任何事件,若它是乌云出现的事件,则它是即将

下雨的事件。

规则(1)显然要依赖相应的全称条件句(2)的可信赖度。因此要估值形如规则(1)的推理可靠度,就必须首先估值形如相应的陈述句(2)存在多大归纳支持。Cohen 认为,必须首先建立归纳支持的理论,来阐述估值归纳支持的规范方法和关于归纳支持的陈述句的逻辑结构。

2. 归纳支持理论的语义基础——相干变量法

为了给归纳支持理论奠定符合科学实践的语义基础,Cohen 提出了相干变量法(the method of relevant variables)。对一科学的特定领域,相干变量法可以用一种概括形式重建如下:

(1) 给定一范畴 $C = C_1 \cup C_2$,其中 C_1 和 C_2 分别由 m 元谓式组成,C_1 中的谓式互不相容(如天鹅、乌鸦),C_2 的谓式为目标谓式(如白的、黑的)。C_1 中任一元与 C_2 中任一元在语言中相容。给定一范畴 C,C 的可检验全称概括条件句 H 构造如下:H 的前件由 C_1 中的元素和个体变元组成的原子公式或多个这样的原子公式的合取构成,H 的后件是 C_2 中若干个元素和个体变元组成的原子公式对真值联结词的布尔组合,例如对 $R_1, R_2 \in C_1, S_1, S_2 \in C_2, (x)(y)(R_1(x) \wedge R_2(y)) \rightarrow (\neg S_1(x,y) \rightarrow S_2(x,y))$ 是可检验概括句。由 C 范畴用以上方式构成的所有可检验的全称条件句的集合记为 CG。

(2) 相对 C,一情境 V 称为相干情境,若它不用 C 的词汇来描述,且它的出现至少证伪 CG 的一元素。例如"在澳大利亚"就能证伪"所有天鹅都是白的"。

(3) 任意互相并列、互不相容的相干情境集构成 C 的一相干变量。包含在一个相干变量中的相干情境称为这个相干变量的变素。例如,相对"所有渡鸦都是黑的",{春、夏、秋、冬},{在交配季节、不在交配季节}都是相干变量。若两相干

变量 v_i 和 v_j 至少有一些变素的可能的组合构成第三个相干变量 $v_k(\subseteq v_i \times v_j)$. 这样的 v_k 称为复合相干变量，否则称为非复合的. 相对 C,此外没有别的相干变量.

(4) C 的相干变量序列可以用下列三个概念定义. 一相干变量称为"极大的"，若它是在并列且不相容的关系下封闭的相干情境集. 一相干变量称为"非穷举的"，若存在某个与该相干变量的任何变素都不相容的不相干情境. 一非复合相干变量称为"比别的相干变量更重要"，若它在 C 的范围内具有更大的证伪潜力.

C 的一相干变量称为 C 的相干变量序列的基元，若它是非复合的、极大的且非穷举的，并且在重要性上不等于其他相干变量. 因此,对于 C，相干变量的一公共序列 (common series)定义为以重要性相对递减的基元进行排序的 C 的相干变量集.

(5) 任意 $H \in \mathrm{CG}$ 的相干变量序列是 C 的公共相干变量序列,且使得该序列的第一个相干变量 V_1 把 H 的前件描述的谓式作为它的唯一元. 例如,若 $(x)(R(x) \rightarrow S(x)) \in \mathrm{CG}$,则 $V_1 = \{R\}$.

(6) 任意 $H \in \mathrm{CG}$ 的归纳可信赖性由一个一元支持函数 $S_C(\cdot)$ 来分级(为了简洁，以后在不易混淆的地方省略脚码 C,但 $S(\cdot)$ 仍然是相对某个范畴而言的).这个函数根据 H 的相干变量序列把 H 映射到 i/n 上(其中 i 和 n 皆正整数且 $i \leqslant n$). 具体说,相对 $H \in \mathrm{CG}$ 的相干变量序列:$V_1, V_2, \cdots, V_n$, $S(H) \geqslant i/n \Longleftrightarrow$ 对所有 $x \in \mathrm{RV}^i$, H 抗拒了 x 的证伪,即在 x 描述的正常情况下没有发现 H 的反例，这里 $\mathrm{RV}^i = \{x: x \in V_1 \times \cdots \times V_i$ 且 x 在物理上可能，其中 $V_1, \cdots, V_i$ 是 H 的相干变量序列的前 i 个相干变量$\}$.

(7) 上述(6)给出的是形如 $S(\cdot) \gtreqqless i/n$ 的陈述句的真

值条件。为了给出陈述句的正当性条件（justification-conditions），即估值 $H \in CG$ 归纳可信赖性的方法，我们定义一典范检验的序列 $t_1, t_2, \cdots, t_n$。给定 $H \in CG$ 的相干变量序列：$V_1, V_2, \cdots, V_n$，就有一相应的典范检验序列 $t_1, t_2, \cdots, t_n$，其中 t_i 由实验背景（在检验 H 时所做的一系列不同的实验中所有不变因素的总和）和 RV^i 构成。对于 $H \in CG$，在进行 t_i 时，对任意 $x \in RV^i$，H 的前件条件的各特例得到考察。若 E 报道，这些前件对 t_i（或对 $t_{i+1}, i+1 \leqslant n$）成立，并且 H 的后件被满足，则 E 为真规范我们断定 $S(H) \geqslant i/n$。若 E 报道在 $t_1, t_2, \cdots$，或 t_i 的一个场合，H 的后件未被满足，则 E 为真就规范我们断定 $S(H) < i/n$。这样，对 H 的归纳证据支持，就是对它的可信赖性给于大于零的等级。

（8）假设检验结果是可重复的（当然是在允许的精度内），若从 E 为真，用上述方法既推出 $S(H) \geqslant i/n$ 又推出 $S(H) < i/n$，则一定存在一个没有排入序列的隐藏的相干变量，它在 E 报道的检验中一直起作用，因此我们需要修正 H 的相干变量序列，或者修正假设的 C 范畴，即引入新的科学概念。

（9）若给定 E 为真，则二元支持函数 $S(\cdots, -)$ 能使我们有权对 H 的可信赖性进行分级。E 报道 H 在 $t_1, \cdots, t_n$ 中通过了 $t_1, \cdots, t_i \Longleftrightarrow S(H, E) \geqslant i/n$。由此可见，一元和二元支持函数不是测度函数而是分级函数。

（10）对适当丰富的语言，可信赖性分级的方法和估值证据支持的方法，可以扩充到检验关于因果联系或定量相关的假说。这时规定 V_1 由 CG 中所有句子的前件提到的所有变素组成，并且构造相应的更复杂的检验（[13], pp. 73—75）。这样，当 $t_1, t_2, \cdots, t_n$ 对应 Mill 的契合法时，t_1 能对应差异法（在因果假设的情况下）或共变法（在定量相关假设的情况

下).

(11) 我们可以进一步扩张 $C_1^* = C_1 \cup$ {谓述相干变量的变素的术语}, 从而使C扩张到 C^*,CG 扩张到 CG^*. 这样对任一 $H \in CG$, 我们可以找到一适当的 $H^* \in CG^*$使得当H未通过 t_i时,把证伪H的变素的否定插入H的前件,或者把未证伪H的变素插入H的前件, 从而得到 H^*. 显然 H^*能抗拒对H的证伪. 例如对 $(x)(R(x) \to S(x)) \in CG$, 在 t_2 中恰被 V_2^3(V_2^3 是H的相干变量序列中第二个相干变量的第三个变素)和 V_1 的某组合证伪, 则下列扩张形式至少有第二级归纳支持

$$(x)(R(x) \wedge V_2^1 \to S(x)),\ (x)(R(x) \wedge \neg V_2^3 \to S(x)).$$

(12) 若两全称条件句属于不同的范畴(也可说两科学假说属于不同的研究领域),一般说, 它们的归纳支持等级是无法比较的,即对于不同的范畴,它们的支持函数一般是互相不可公度的.

3. 相干变量法和可能世界语义模型

Cohen 认为他的系统可以用可能世界语义模型来解释. 对具有 n 个相干变量的 V_1 和目标谓式的任一特定的范畴,一逻辑可能世界 w_1(相对 CG)受到另一可能世界 w_2 的齐一性制约(若用R表示这种制约关系, 则 Rw_1w_2,Cohen 也称R是归纳通达关系)⇔(1)对任意 $H \in CG$, 若H在 w_2 上为真, 且被例证, 则H在 w_1 上真,不管H在 w_1 上空洞地真(即H的前件假) 还是被例证. 且(2) 若H在 w_2 上空洞地真, 则H也在 w_1 上空洞地真. 一逻辑可能世界w称为物理可能世界⇔ Rww_e (其中 w_e 表示现实世界). 相干变量序列无需在此重新定义, 但一情境C(不用 V_1 和目标谓式的术语描述) 称为相干⇔存在 $H \in CG$,使得H在C和H的前件被联合例证的某个物理可能世界上为假, 且H至少在H的前件被例证而C不

被例证的某个物理可能世界上真. 一物理可能世界 w 称为 $t_{1/2}$ 世界⇔任意 $x\in RV^1$ 在 w 上被例证. 一物理可能世界 w 称为 $t_{2/3}$ 世界⇔对每一 $t_{1/2}$, 都有 $Rt_{1/2}w$, 且任意 $x\in RV^2$ 在 w 上被例证, …. 现实世界 w_e 和每一 t_n 世界包含大量的事件,使得 V_1 的变素和其他相干变量的 0 个或多个变素的每一物理上可能的组合在这些世界上被例证. 并且对所有 i 和 j $(i\leqslant j)$,都有 Rt_it_j 且 Rt_jt_e.

因此对任意 $H\in CG$ 或H是 CG 中有限元素的合取, $\Box^iH(i<e)$真 $\Leftrightarrow$ H在所有 t_i 上真(不管是空洞地真还是被例证). $\Box^eH$ 真 $\Leftrightarrow$ H不仅在所有 $t_i(i<e)$ 上真, 而且还在现实世界上真. $\Box^dH$ 真 $\Leftrightarrow$ H在所有的逻辑可能世界上真. 对任意 $H\notin CG$ 或不是 CG 中有限元素的合取, $\Box^iH$ 真⇔存在一公式 I 使得$\Box^iI$ 真且在所有 I 为真的逻辑可能世界上H真.

Cohen 认为"……受……的齐一性制约"构成了归纳可通达或归纳可知关系,这关系在某个模型的某些世界之间成立,而在另一些世界之间不成立. 因为这个关系是传递, 自返且不对称的, 所以对上述非形式的表述给予适当的形式化就是对 C. I. Lewis 的模态系统 S_4 的概括.

因此 $H\in CG$ 的先验归纳概率(等于归纳可信赖性)可以看作是H的归纳适域 (range) 的排级,而 Carnap 的关于一命题的先验数学概率可以看作是该命题的逻辑适域的测度, H的归纳适域由H成立的那个最完备的物理可能世界 (最高等级的相干事件世界) 来排级. 而一命题的逻辑适域由指派给此命题成立的各逻辑可能世界的赋值的总和来测度. 因此, Cohen 认为他的逻辑着眼于可能世界在物理上确定的质,而 Carnap 等人的逻辑着眼于这些世界在逻辑-语言上确定的量. 这就是两种逻辑为何最终不同的根本原因. Cohen

认为他自己的逻辑对自然科学中的实验推理提供了一种更好的表述。

4. 归纳支持与数学概率的不可公度性(incommensurability)

根据上述归纳支持理论,显然可以得出一重要的推论:既使当 E 和 H(全称条件句)矛盾时,也可能有 $S(H,E)>0$,这是因为 E 可以报道 H 通过 t_i 但没有通过 t_{i+1} (后面的公理化系统也能证明该结果)。从这里至少可以得出,二元归纳支持函子和二元数学概率函子的逻辑语法不同,因为对于后者,若 E 和 H 矛盾,则 $P_M(H,E)=0$。

因此可以证明上述归纳支持的任何概念不仅不是数学概率本身,而且不是数学概率的任何函数,亦即归纳支持概念的逻辑语法不能以任何方式归约为概率的数学演算。下面就是 Cohen 的证明。

假设 $S(H,E)$的值是关于 H 和 E 的数学概率的函数,其中 H 和 E 在数学概率演算中是可表达的。易证,对任意 H 和 E,当 $1>P_M(E)>0$ 时,我们有(i)$P_M(\neg H,E)$和 $P_M(\neg H)$ 分别是 $P_M(H,E)$ 和 $P_M(H)$ 的函数;$P_M(H\wedge E,H)$ 是 $P_M(E,H)$的函数,$P_M(H\wedge E,E)$是 $P_M(H,E)$ 的函数,$P_M(H\wedge E)$是 $P_M(H)$ 和 $P_M(E,H)$ 的函数。(ii)根据概率的数学演算,当 $P_M(E)>0$ 时,$P_M(H,E)=P_M(E,H)\cdot P_M(H)/P_M(E)$。(iii)$P_M(H,H\wedge E)=P_M(E,H\wedge E)=1$。因为 $\neg$ 和 $\wedge$ 是功能完备的,所以对任意 H,E,若 $1>P_M(E)>0$,则相对 $S(H,E)$,只需考虑 $P_M(E,H)$,$P_M(H)$ 和 $P_M(E)$。

现考虑两逻辑上互相独立的全称条件句 H 和 H',使得 $P_M(H)=P_M(H')$。令 E 报道检验 t_i(在 t_i 下 H 被证伪,H' 却通过)和 t_j($j>i$,在 t_j 下 H' 被证伪),则 $1>P_M(E)>0$,且

$S(H', E) > S(H, E)$. 因为E和 H, H' 都矛盾, 并且 $P_M(H) = P_M(H')$, 所以 $P_M(E, H) = P_M(E, H') = 0$.因此$H$和 H' 的归纳支持等级不同, 而我们所考虑的每一对数学概率相同.这就得出,对任意H和 E,当 $1 > P_M(E) > 0$ 时, $S(H,E)$不是 $P_M(E,H)$,$P_M(H)$ 和 $P_M(E)$ 的函数. 因此归纳支持和数学概率不可公度.

§3 归纳支持分级的公理化系统

Cohen 的系统是用语法的元语言建立起来的. 相对每一特定种类的对象语言,这种元语言统一描述形成规则、公理和推演规则. 对象语言用于不同的科学研究领域,对这些领域,不同序列的相干变量和与之对应的归纳支持函数和归纳概率函数(后者在§4 讨论)都是恰当的. 因此,各种对象语言在它们非逻辑术语的数量和本性上, 在它们能够表达的归纳支持或归纳概率的分级的数量上, 可以是互相不同的. 但是它们全都具有可以用单一的语法的元语言来描述的相同的逻辑术语,相同的逻辑原则.

现在我们表述 Cohen 系统.

1. 元语言符号

(1) $x, y, z, x_1, y_1, z_1, \cdots$表示对象语言中的个体符号.若它们在公式中的出现不是约束的,则假设它们什么也没表示.

(2) Q, R, S, Q_1, $\cdots$表示一阶m元谓词符号. 若它们在公式中的出现不是约束的, 则假设它们描述了科学研究领域的一个特性 (就象如§2 中提到的C范畴中的谓式或目标谓式).

(3) $\neg, \wedge, \exists$,(,), 是自指的, 其中$\neg, \wedge$在对象语言中分别具有它们通常的解释,作为否定和合取的记号,$\exists$ 表示存

在量词,(,)是括号.

(4) □表示模态算子常元和变元. 在任何一种对象语言中,模态算子常元是根据不同的数码上标来互相区别. 在元语言中,$\Box^d$表示对象语言的一特指的模态算子,即具有最高上标的模态算子. $\Box^d H$ 可以解释为: H逻辑真或分析真,也可以理解为H具有逻辑必然性. $\Box^e$ 表示上标 e 最接近 d 的模态算子,也可以解释为: 命题H具有完全的归纳可信赖性. 假定H的相干变量的数目仅有 e 个且H通过 t_e, 则称H具有完全的归纳可依赖性. $\Box^i H(i < e)$ 可以解释为命题H被赋于某个低于完全归纳可信赖性的等级.

Cohen 认为: 当从R到S的推理的规则越来越接近自然规律(有时 Cohen 把$\Box^e H$ 看作H是一自然规律)时,根据 R, S 的归纳概率也越来越高,而归纳概率就是对自然必然性的分级. 这就是为何用一种模态逻辑的形式系统来表示是恰当的. 相应地,对于这个形式系统,一可能世界的模型可以看作是表示与相干变量法一致的受控实验的结构([14],p.65).

在特定的研究领域,每一恰当的估值归纳支持方法的工作假说,都应规定一确定的有穷的相干变量序列. 但是从本体论的观点来说,总存在逻辑上的可能,使相干变量的数目是可数无穷的. 因此在描述的各对象语言中,我们除了假设存在某个只能表达不同的有穷数目的归纳支持等级的对象语言外,还假设存在某个能表达可数无穷个不同的归纳支持等级的对象语言. 此外可以令 $d = 2, e = 1$, 初始模态算子具有的最低上标为 1/2, 并且假定,若存在上标为 $r/(r+1)$ 的模态算子常元,则还存在一上标为 $(r+1)/(r+2)$ 的模态算子常元,直到我们所需要的模态算子常元序列的最高上标低于 e. 若这些上标小于 1 的模态算子常元根据上标数的大小排列起来,就有 $\Box^{1/2}$, $\Box^{2/3}$, $\Box^{3/4}$, …,则 $\Box^{1/2}H$ 表示H至少

存在第一级归纳可信赖性，$\Box^{2/3}H$ 表示H至少存在第二级归纳可信赖性，以此类推，因此归纳可信赖等级可以充分地由这样的上标的分子来表示。相干变量的数目 n 和正支持等级（大于 0）由（除 d 和 e 之外的）最高上标的分母和分子来表示，若存在这样一个上标，否则就用∞表示。因此这样的分数上标所起的作用只是作为一种方便的概念装置。

模态算子变元用 $\Box^X,\Box^Y,\cdots$表示。若对象语言只有有穷个模态算子常元，就可以完全省略模态算子变元的概念。但是，因为某些对象语言被认为具有可数无穷个模态算子常元，所以统一地构造能处理所有这样的对象语言的逻辑，并且在它们每一个中引入模态算子变元将是方便的。

为了方便，在本系统中统一使用 u 和 w 来表示个体符号、谓词符号和模态算子变元，除非另作规定。例如 $(u)H\rightarrow H$ 可以统一表示$(\Box^X)H\rightarrow H$，$(x)H\rightarrow H$ 和$(R)H\rightarrow H$。

2. 形成规则

(1) $R(x_1,\cdots,x_m)$ 是原子公式，其中R是m元谓词。

(2) 若 A,B 是公式，则 $\neg A, A\wedge B$，$\Box^i A,(\exists x)A$ 和 $(\exists R)A$ 都是公式。

(3) 若A是公式，B只是在下列方面与A不同：A中有$\Box^i$出现的一个或多个地方，B中有$\Box^X$出现，则 $(\exists\Box^X)B$ 也是公式。

今后我们用 E,F,G,I,J 等来表示公式。

3. 定义

除了通常对$\vee,\rightarrow,\longleftrightarrow$的定义，还有

$$(u)H\triangleq\neg(\exists u)\neg H, \tag{6.1}$$

$$(H\rightarrow^i I)\triangleq\Box^i(H\rightarrow I), \tag{6.2}$$

$$(H\longleftrightarrow^i I)\triangleq((H\rightarrow^i I)\wedge(I\rightarrow^i H)), \tag{6.3}$$

$$\Diamond^i H\triangleq\neg\Box^i\neg H, \tag{6.4}$$

$$\Box^0 H \triangleq \Box^{\frac{1}{2}} H \vee \neg \Box^{\frac{1}{2}} H. \tag{6.5}$$

今后用$\Box^i$和$\Box^j$统一地表示任何初始模态算子常元或$\Box^0$。一元和二元归纳支持函数 $S(\cdot)$，$S(\cdots,—)$ 可以根据上下文自指地引入。在存在 n 个等级的归纳可信赖性和 $0 \leqslant i \leqslant e$ 的条件下，根据上下文把 $S(\cdot)$ 定义为

$$S(H) \geqslant i \triangleq \Box^i H, \tag{6.6}$$

$$S(H) < i \triangleq \neg(S(H) \geqslant i), \tag{6.7}$$

$$S(H) \leqslant i \triangleq \neg \Box^j H, \tag{6.8}$$

其中当 $i = 0$ 时，$j = 1/2$；当 $i(<e) = r/(r+1)$ 时，$j = (r+1)/(r+2)$；当 $i = e$ 时，$j = d$（即通过否认运用更高上标的模态算子常元来断定H的可信赖性不高于某个等级）。

$$S(H) \geqslant S(I) \triangleq (\Box^X)(\Box^X I \rightarrow \Box^X H). \tag{6.9}$$

在具有有穷数目 n 的模态算子常元的对象语言中可以用下列有穷合取来置换(6.9)：

$$(S(I) \geqslant 0 \rightarrow S(H) \geqslant 0) \wedge (S(I) \geqslant 1/2 \rightarrow S(H) \geqslant 1/2) \wedge \cdots \wedge (S(I) \geqslant (n-1)/n \rightarrow S(H) \geqslant (n-1)/n) \wedge (S(I) \geqslant e \rightarrow S(H) \geqslant e),$$

$$S(H,E) \geqslant i \triangleq (E \rightarrow {}^e\Box^i H), \tag{6.10}$$

$$S(H,E) \leqslant i \triangleq \neg(E \rightarrow {}^e\Box^j H), \tag{6.11}$$

其中 i 和 j 的关系如(6.8)。

$$S(H,E) \geqslant S(I,F) \triangleq (\Box^X)((F \rightarrow {}^e\Box^X I) \rightarrow (E \rightarrow {}^e\Box^X H)). \tag{6.12}$$

一元定理是一断定，它具有这样的形式："H 是可证的"，可以缩写为"$\vdash H$"。

4. 公理模式和推理规则

注意：对正规模态逻辑 S_4 的概括，我们需要（6.13)—(6.15)，(6.19)，而(6.23)—(6.28)是和标准量化原则一致的。

若H是重言式及其代入特例,则$\vdash H$, (6.13)

若$\vdash H$且$\vdash H \to I$,则$\vdash I$, (6.14)

若$\vdash H$,则$\vdash \Box^d H$, (6.15)

若$\vdash H$,则$\vdash \Diamond^i H$, 其中 $0 < i < e$, (6.16)

$\vdash (H \to {}^i I) \to (\Box^i H \to {}^i \Box^i I)$, 其中 $0 < i$, (6.17)

$\vdash (H \to {}^i I) \to (\Box^i H \to \Box^i I)$, 其中 $0 < i < e$, (6.18)

$\vdash \Box^i H \to \Box^j H$, 其中 $i < j$, (6.19)

$\vdash \Box^e H \to H$, (6.20)

$\vdash \Box^i H \to \Box^i (x) H$, 其中 $0 < i$, (6.21)

$\vdash \Box^d H \to (R) H$, (6.22)

$\vdash H \to (u) H$,其中u在H中不自由出现, (6.23)

$\vdash (u) A \to H$, (6.24)

其中,若(1)H与A相同,除了A有u自由出现之处,H有y或R或$\Box^i (0 < i \leqslant e)$自由出现,并且(2)若在$A$的形如$\Box^i G$或$\Box^X G$的子公式中有$x$的任意自由出现,则$H$是从$A$用$y$处处代入上述$x$得到的公式使得$y$是自由出现且$y$不同于$(x)A$中自由出现的个体符号.

$$\vdash (u)(A \to B) \to ((u) A \to (u) B), \qquad (6.25)$$

其中 A, B 是公式(可以包含自由出现的$\Box^X$,$\Box^Y$,…).

若$\vdash H$,则$\vdash (x) I$, (6.26)

其中I与H相同,除了在H有y自由出现之处,I有x自由出现.

若$\vdash H$,则$\vdash (R) I$, (6.27)

其中I与H相同,除了在H有S自由出现之处,H有R自由出现.

若$\vdash H$,且H中$\Box^i$出现$(0 < i \leqslant e)$,并且若对H中$\Box^i$出现的某些地方和对于所有i使得$0 < j \leqslant e$,有$\vdash G$,其中G和H相同,除了在上述地方G有$\Box^j$自由出现,则

$$\vdash (\Box^x)A, \tag{6.28}$$

其中A与H相同，除了在上述地方有$\Box^x$自由出现.

Cohen 在[15]和[13]证明了几百个定理，我们在此只举几个我们感兴趣的定理.

定理 6.1 一元归纳支持的特殊否定原则

$$S(H) > 0 \to S(\neg H) = 0.$$

注意该定理与数学概率演算的相应定理的不同之处.

定理 6.2 一元归纳支持的特殊合取原则

$$S(H) > 0 \wedge S(I) > 0 \to S(H \wedge I) > 0.$$

定理 6.3 二元归纳支持的特殊否定原则

$$\Diamond^e E \to (S(H,E) > 0 \to S(\neg H, E) = 0).$$

此原则可以与 Carnap 的定理 3.13(8): $c(\neg h, e) = 1 - c(h, e)$ 相比较.

定理 6.4 二元归纳支持的特殊合取原则

$$S(H,E) > 0 \wedge S(I,E) > 0 \to S(H \wedge I, E) > 0.$$

定理 6.5 二元归纳支持的一般合取原则

$$S(H,E) \geqslant S(I,E) \to S(H \wedge I, E) = S(I,E).$$

定理 6.6 $((E \longleftrightarrow {}^i F) \wedge (H \longleftrightarrow {}^j I)) \to S(H, E) = S(I,F)$,其中 $i \geqslant e \leqslant j$.

§4 归纳概率分级的公理化系统

1. 归纳概率与数学概率之间某些逻辑上的差别

二元归纳概率函数 P_I 和二元数学概率函数 P_M在逻辑语法上存在的一个重要差别是，前者在变目位置上的命题在逆否变换下函数值保持不变，而后者则不然(在这点上，归纳支持的逻辑语法与数学概率函数的逻辑语法相同).因为 $(x)(R \to S)$ 逻辑等值 $(x)(\neg S \to \neg R)$，并且逻辑等值的命题有相

同的归纳可信赖等级，所以对归纳概率函数，相应的概括句的前后件的可信赖性在逆否变换下不变，即对任意 S 和 R，$P_I(S,R)=P_I(\neg R,\neg S)$（参见后面定理 6.10），而 $P_M(S,R)$ 不一定等于 $P_M(\neg R,\neg S)$，除非 $P_M(S,R)=1$（参见[2],435）. 对支持函数，$S(H,I)$不一定等于 $S(\neg I,\neg H)$，除非$\vdash(H\overset{i}{\longleftrightarrow}\neg I)$，其中 $i\geqslant e$，可参见定理 6.6.

归纳概率和数学概率的另一个重要差别在于一元和二元函数之间的关系. 在归纳支持系统中，我们有下列定理:

$$S((x)(R(x)\rightarrow S(x)))\geqslant S((x)\neg R(x))$$

和

$$S((x)(R(x)\rightarrow S(x)))\geqslant S((x)S(x)).$$

于是根据§2 的有关内容，就有 $P_I(S,R)\geqslant P_I(\neg R)$ 和 $P_I(S,R)\geqslant P_I(S)$，其中 $R\rightarrow S$ 是 $(x)R(x)\rightarrow S(x)$ 的代入特例，这类似于所谓的真值函数蕴涵悖论. 但是，虽然归纳概率的这些原则不同于数学概率或归纳支持的相应的原则，但它们却不是悖论. 因为，(1)，根据§2 的 2. 的(1)的定义，形如 $(x)R(x)$ 的概括句不是可检验的，因此这样的概括句不能用通常的方式得到归纳支持，只有在相当特殊的情况中，由于它和别的概括句的逻辑关系，才能获得归纳支持. 因此在正常情况下，我们应该有 $P_I(\neg R)=0$，且我们不能从 $P_I(S,R)\geqslant P_I(\neg R)$得出 $P_I(S,R)$ 的值. (2)，从 $P_I(S)=S(S)$，$P_I(S,R)\geqslant P_I(S)$ 推出，若 S 在任何场合都有一定的归纳支持等级（这我们能从适当的实验和观察中得到），则 S 根据任何相关的证据，至少具有这种等级的归纳概率. 然而，虽然当 $P_M(R)>0$ 时，$P_M(S)=0$ 蕴涵 $P_M(S,R)=0$，但是对于归纳概率（或归纳支持）却没有相应的原则存在. 这是因为 $P_I(S)=0$ 蕴涵的只是不存在先验的归纳理由来相信 S，而不是蕴涵存在相信$\neg S$ 的先验的归纳理由. 若我

们需要一个特定的理由，依靠归纳根据来期望 S，则我们完全可以有理由有 $P_I(S)=0$ 且 $P_I(S,R)>0$。

第三个重要区别是归纳概率和数学概率的合取原则，这是由于归纳支持和数学概率的合取原则存在巨大的差别。归纳概率的合取原则直接来自归纳支持的合取原则，形式上也与它一致：

若 $P_I(S_1,R)\geqslant P_I(S_2,R)$，

则 $$P_I(S_1\wedge S_2,R)=P_I(S_2,R). \qquad (6.29)$$

第四个重要区别是否定原则。归纳概率的否定原则是非补余性的。假定 S_1 和 S_2 是互相矛盾的，且 $P_I(S_1,R)\geqslant P_I(S_2,R)\geqslant 0$，根据(6.29)，对自相矛盾的合取式，$R$ 给出某个非零的概率等级。因为这样的合取的否定是重言式，所以根据逆否变换，存在 R 为假的先验概率：

若 $P_I(S,R)\geqslant i/n$，且 $P_I(\neg S,R)\geqslant i/n$，则
$$P_I(\neg R)\geqslant i/n. \qquad (6.30)$$

令 $i=1$ 和逆换，我们就得到归纳概率的否定原则：

若 $P_I(S,R)>0$ 且 $P_I(\neg R)=0$，则 $P_I(\neg S,R)=0$。

这非常类似归纳支持的否定原则（定理 6.3），也只有该原则才说明归纳概率是作为不完全系统的可证性的分级。

大概有人会因为根据某证据，归纳概率给自我矛盾的命题以非 0 值而提出异议，认为这太不符合直观。Cohen 认为这种反对意见是完全错误。因为，首先自我矛盾的命题的先验概率的确为 0．但在归纳逻辑中，为 0 的先验概率不能硬性规定为 0 的后验概率，因为归纳概率函数只是对证据的权作出评价（[13]，pp. 36—39）。第二，(6.30)表明的只是二元归纳概率逻辑中通过归谬法论证的逻辑原则的概括形式，就象先假设存在最大素数，然后推出矛盾，从而证明不可能存在最大素数那样。

此外，Cohen 在[13]中还证明了 $P_I(S, R)$ 不可能是 $P_M(R,S)$，$P_M(S)$ 和 $P_M(R)$ 的某个函数，从而证明了归纳概率和数学概率之间也是不可公度的。

2. 归纳概率分级的公理化系统

据前述，归纳概率概念是依赖于归纳支持概念的，所以 Cohen 建立的公理化的归纳概率系统就是在 §2 给出的系统中做如下增补：为了引入二元概率函子，用 $\alpha(\alpha_1,\alpha_2,\cdots)$ 和 $\beta(\beta_1,\beta_2,\cdots)$来表示不含□和量词的公式。在下面的定义中，它们表示，任意个体符号在 $\alpha_1,\alpha_2,\cdots$有一次或多次出现 ⇔ 它在 $\beta_1,\beta_2,\cdots$也有一次或多次出现，

$$P_I(\alpha,\beta) \geqslant i \triangleq S(\beta \to \alpha) \geqslant i, \qquad (6.31)$$

P_I 是归纳概率算子，为了简略以后省去脚码。另外

$$P(\alpha) \geqslant i \triangleq (P(\alpha,\alpha \vee \neg\alpha) \geqslant i). \qquad (6.32)$$

一元和二元归纳概率函数的其它定义类似前面关于归纳支持的定义，只是对二元归纳概率函数，要注意变目取值应是 α，β，$\cdots$。

下面我们列举几个重要的定理：

定理 6.7 一元否定原则

$$P(\alpha) > 0 \to P(\neg\alpha) = 0.$$

定理 6.8 一元后承原则

$$(\alpha_1 \to \alpha_2) \to P(\alpha_2) \geqslant P(\alpha_1).$$

定理 6.9 一元合取原则

$$P(\alpha_2) \geqslant P(\alpha_1) \to P(\alpha_1 \wedge \alpha_2) = P(\alpha_1).$$

定理 6.10 逆否原则

$$P(\alpha,\beta) = P(\neg\beta, \neg\alpha).$$

定理 6.11 二元合取原则

$$P(\alpha_1,\beta) \geqslant P(\alpha_2,\beta) \to P(\alpha_1 \wedge \alpha_2,\beta) = P(\alpha_2,\beta).$$

由此可见 Cohen 的归纳概率语法系统不过是他的归纳

支持语法系统的保守扩张。

可以证明若标准二阶量化理论是一致的（这一点在 A. Church "Introduction to Mathematical Logic", (1956), Vol. I 已证明），则 Cohen 的系统也一致。为了把上述系统中的公式解释成标准量化理论的公式，只要在每一公式中，用 $(R_1)(R_2)\cdots(R_n)(x_1)(x_2)\cdots(x_m)A$ 去置换 $\Box^i A$ 或 $\Box^X A$ 的每一次出现（这里的 $R_1, R_2, \cdots, R_n, x_1, x_2, \cdots, x_m$ 是 A 中所有自由出现的谓词和个体变元），用 $(\exists y)A$ 来置换 $(\exists\Box^X)A$ 的每一次出现（这里的 y 在 A 中不自由出现）。这样，我们有 (6.13)，(6.14)，(6.23)—(6.27)作为公理或规则就足够了，因为其余的都可以从上述公理或规则推出。

总之，Cohen 对归纳逻辑的建立作了新的探索，建立了旨在刻划消除归纳法（Carnap, Reichenbach 等人的归纳逻辑主要刻划枚举归纳法）的模态归纳逻辑，这样的工作是有益的。但整个来说，他的工作还十分粗糙，主要缺陷表现在：他的作为语义模型的相干变量法既繁琐又含混，不能令人满意地解释他的公理化语法系统。由此使人无从考虑完全性定理。看来建立简单而又解释力强的形式语义模型，以至语义学，并且相应地建立反映科学实验中的归纳推理形式（主要是消除归纳法）的形式语法系统仍是归纳逻辑学家今后要做的工作。

第七章　模态归纳逻辑(下)

——Burks 的因果陈述句逻辑

A. W. Burks 的因果陈述句逻辑也是一种模态归纳逻辑。和 Cohen 的逻辑类似，因果陈述句逻辑也用一种概括了的模态逻辑系统作为自己的语法系统，且对该系统的基本概念，即必然算子作了归纳意义的解释。我们知道因果关系和消除归纳法有密切的联系，F. Bacon, J. S. Mill 和 D. Hume 都看到了这一点。Bacon 和 Mill 通过对因果关系的分析建立了古典归纳逻辑（其核心是 Mill 的求因果关系五法：契合法、差异法、契合差异并用法、共变法和剩余法），而 Hume 则从否定性的方面对因果关系和归纳推理作了深入的分析，他的分析如此深刻，以致后人把如此产生的归纳问题称为休谟问题。Burks 继承了 Bacon 和 Mill 的传统，运用模态逻辑、概率演算和其他现代知识重新考察了因果关系和归纳法的联系，建立了因果陈述句逻辑，为建立合理刻划因果关系的模态归纳逻辑做了有益的探索。

本章§1，我们介绍和分析 Burks 的因果陈述句逻辑的语法系统。§2，讨论 Burks 给出的两个初步的解释：逻辑解释和具体解释，在 §3，我们介绍和讨论 Bursk 的归纳理论。在那里 Burks 力图把因果必然性和归纳概率联系起来，并且探讨了归纳概率和因果陈述句逻辑的联系。

Burks 的模态归纳逻辑内容十分丰富，我们在此受篇幅的限制，介绍和讨论都非常概要，有兴趣的读者请参见[16]，

§1 因果陈述句逻辑的语法系统

什么是因果陈述句？我们先来看一个例句：

若这只戒指（r）是金的(G)且把它置于王水(A)，则它将溶解(D)。 (7.1)

显然（7.1）是化学中的真陈述句，因此是一个经验真的陈述句。若把(7.1)看作是一个实质蕴涵句，则(7.1)可以形式化为“$G(r)\wedge A(r)\to D(r)$”。因为“$\neg A(r)\to(G(r)\wedge A(r)\to D(r))$”是普遍有效式，所以“$G(r)\wedge A(r)\to D(r)$”可从该戒指不放入王水的假设（“$\neg A(r)$”）中推出，但在实际过程中（例如在化学中）(7.1)显然不能单从“$\neg A(r)$”中得到，因此(7.1)不能看作是实质蕴涵句。因为(7.1)不是逻辑真的，所以把(7.1)看作是严格蕴涵句也是不合适的。可见，(7.1)是一种新型的蕴涵句，Burks 称它为因果蕴涵句，由于这种蕴涵句与因果必然性相关，因此 Burks 认为有必要建立一种逻辑来研究它。

Burks 的因果陈述句逻辑的初始符号分为两类：常元和变元。逻辑常元是真值联结词“$\neg$”和“$\vee$”，左右括号“(”，“)”（用于隔断和量化），和模态符号“$\Box$”，“$\Box^c$”。变元有下列三类：

陈述句变元：$A, B, Q, A_1, B_1, \cdots$；

个体变元：$a, b, x, y, x_1, y_1, \cdots$；

谓词变元：$A, B, C, A_1, A_2, \cdots$。

公式的形成规则如下：

(1) 陈述句变元是公式。若 $x_1, \cdots, x_n$ 是个体变元，A 是谓词变元，则 $A(x_1, \cdots, x_n)$ 是公式。

(2) 若 φ，ψ 是公式且 x 是个体变元，则 $\neg\varphi$，$\varphi\vee\psi$，$(x)\varphi$ 是公式。若 φ 是不包含$\Box$和 $\Box^c$ 出现的公式，则$\Box\varphi$和$\Box^c\varphi$是公式。

从直观上说，用“$\Box\varphi$”表示φ是逻辑必然的，用“$\Box^c\varphi$”表示φ是因果必然的，所以上述(7.1)可以表示为“$\Box^c(G(r)\wedge A(r)\rightarrow D(r))$”，或简写为“$G(r)\wedge A(r)\underset{c}{\rightarrow}D(r)$”。由于Burks认为在实际的因果关系和归纳法中叠加的模态词不起什么作用，所以可以从逻辑上排除这种情况，这一点与Cohen的模态归纳逻辑不同。此外，还须注意，这里Burks没有给出谓词变元和陈述句变元的量化定义(在后面的§2，他才非正式地引入)。

除了通常对 $\wedge,\rightarrow,\leftrightarrow,(\exists x)$ 的定义，Burks的系统还增加以下定义：

$\Diamond\varphi\triangleq\neg\Box\neg\varphi$,

$\Diamond^c\varphi\triangleq\neg\Box^c\neg\varphi$,

$\varphi\underset{c}{\rightarrow}\psi\triangleq\Box^c(\varphi\rightarrow\psi)$,

$\varphi\prec\psi\triangleq\Box(\varphi\rightarrow\psi)$.

Burks的公理化系统由下列公理(模式)和推理规则组成：

公理 7.1一阶谓词演算的普遍有效式及其代入特例。

公理 7.2(模态词的次序)

(1) $\Box\varphi\rightarrow\Box^c\varphi$,

(2) $\Box^c\varphi\rightarrow\varphi$.

公理 7.3(模态词的分配)

(1) $\Box(\varphi\rightarrow\psi)\rightarrow(\Box\varphi\rightarrow\Box\psi)$,

(2) $\Box^c(\varphi\rightarrow\psi)\rightarrow(\Box^c\varphi\rightarrow\Box^c\psi)$.

公理 7.4(模态词与量词的交换律)

(1) $(x)\Box\varphi\longleftrightarrow\Box(x)\varphi$,

(2) $(x)\Box^c\varphi\longleftrightarrow\Box^c(x)\varphi$.

规则 7.1(分离规则的概括)

$$\varphi_1,\cdots,\varphi_n,\varphi_1\wedge\cdots\wedge\varphi_n\rightarrow\psi/\psi.$$

规则 7.2 $\varphi/(x)\varphi$.

规则 7.3 若φ中不包含$\Box$(或$\Box^c$)的出现，则 $\varphi/\Box\varphi$(或 $\varphi/\Box^c\varphi$).

$\vdash\varphi$和 $\varphi_1,\cdots,\varphi_n\vdash\varphi$的定义如通常. 从上述系统，我们有下列导出规则和定理:

(1) 若 $\varphi_1,\cdots,\varphi_{n-1},\varphi_n\vdash\psi$,则 $\varphi_1,\cdots,\varphi_{n-1}\vdash\varphi_n\to\psi$.

(2) 若 $\varphi_1,\cdots,\varphi_n\vdash\psi$, 则 $\Box\varphi_1,\cdots,\Box\varphi_n\vdash\Box\psi$.

(3) 若 $\varphi_1,\cdots,\varphi_n\vdash\psi$, 则 $\Box^c\varphi_1,\cdots,\Box^c\varphi_n\vdash\Box^c\psi$.

(4) 若 $\varphi\vdash\psi$, 则$\Diamond\varphi\vdash\Diamond\psi$.

(5) 若 $\varphi\vdash\psi$, 则$\Diamond^c\varphi\vdash\Diamond^c\psi$.

(6) $\vdash\Box^c\varphi\wedge\Box^c\psi\longleftrightarrow\Box^c(\varphi\wedge\psi)$.

(7) $\vdash(\varphi\underset{c}{\to}\psi)\wedge(\psi\underset{c}{\to}\theta)\to(\varphi\underset{c}{\to}\theta)$.

(8) $\vdash\neg\Diamond^c\varphi\to(\varphi\underset{c}{\to}\psi)$.

最后一个定理称为因果蕴涵悖论.

§2 对因果陈述句逻辑的语法系统的初步解释

本节 Burks 给予§1建立起来的语法系统两个初步的解释:

1. 抽象解释

抽象解释就是用模态模型 $\mathfrak{A}_{\mathscr{L}}$ 来解释 §1 的语法系统.

考虑语言$\mathscr{L}$,它是在§1 表述的因果陈述句逻辑的基础上通过增加有穷或无穷个个体常元和有穷个任意元(有穷元)的谓词常元而构建成的. 因此前面关于公式的定义(形成规则)应该扩张使得包括这些新符号，而且使得公理和推理规则也能运用于这些新公式上. 在语言$\mathscr{L}$的非逻辑常元和模型 $\mathfrak{A}_{\mathscr{L}}$ 的元素之间有一个一一对应. $\mathscr{L}$ 的每一个体常元命名 $\mathfrak{A}_{\mathscr{L}}$ 的一个体，每一m元的谓词对应 $\mathfrak{A}_{\mathscr{L}}$ 的m元的基本属性. 在不

致混淆的地方，我们把 $\mathfrak{A}_{\mathscr{L}}$ 简记为 $\mathfrak{A}$.

由 $\mathscr{L}$ 的个体常元和谓词常元构成的原子公式称为原子陈述句. 假设这些原子陈述句以某个次序排成一序列 $\langle\varphi_i\rangle$. 我们称 $\bigwedge_i \varphi_i'$ 为基本合取，其中 i 跑遍 $\langle\varphi_i\rangle$ 中所有原子陈述句的下标，φ_i' 是 φ_i 或 $\neg\varphi_i$. 若语言 $\mathscr{L}$ 的个体数有穷，则 $\bigwedge_i \varphi_i'$ 是一有穷句，否则称为无穷句(关于无穷句，我们在最后一章还要用到).

模型 $\mathfrak{A}$ 的世界描述集是 $\mathscr{L}$ 的任意指定的非空基本合取集. 世界描述集的一个非空子集指定为因果可能世界描述集，其中有一元素规定为语言 $\mathscr{L}$ 的现实世界描述. 从直观上说世界描述集就是逻辑可能世界集，其中每一元素自身是无矛盾的，这类似 Carnap 在第三章开头要求其语言满足的独立性. 如果在一个逻辑可能世界上它所涉及的所有因果律成立，则该逻辑可能世界就被称为因果可能世界. 若一个基本合取正好是 $\mathfrak{A}$ 的所有真原子陈述句或原子陈述句的否定的合取，即 $\bigwedge_i\{\varphi_i' : \varphi_i'$ 是原子陈述句或原子陈述句的否定且 φ_i' 在 $\mathfrak{A}$ 中真$\}$，则该基本合取称为现实世界.

下面我们来递归定义"$\mathfrak{A}\models\varphi$"，即 φ 在 $\mathfrak{A}$ 中真：(1)若 $P(a_1,\cdots,a_n)$ 是任意原子陈述句，则 $\mathfrak{A}\models P(a_1,\cdots,a_n)\Longleftrightarrow$ 个体序列 $a_1,\cdots,a_n$ 有属性 P (用经典模型论的术语，即 $\langle a_1,\cdots,a_n\rangle\in P^{\mathfrak{A}}\subseteq A^n$，其中 A 是 $\mathfrak{A}$ 的论域，A^n 表示 n 阶笛卡尔积). (2) $\mathfrak{A}\models\neg\varphi\Longleftrightarrow\mathfrak{A}\not\models\varphi$($\varphi$ 在 $\mathfrak{A}$ 中不真). (3) $\mathfrak{A}\models\varphi\wedge\psi\Longleftrightarrow\mathfrak{A}\models\varphi$ 且 $\mathfrak{A}\models\psi$. (4)令 $a_1,a_2,\cdots$，是 $\mathscr{L}$ 的所有个体常元，$\mathfrak{A}\models(x)\varphi\Longleftrightarrow$ 对所有 a_i，$i=1,2,\cdots$，$\mathfrak{A}\models\varphi(a_i)$，其中 $\varphi(a_i)$ 是用 a_i 代换 $\varphi(x)$ 中所有自由出现的 x 得到的公式. Burks 这里的表述不太严谨，因为 φ 中还可以有别的自由变元. (5) $\mathfrak{A}\models\Box\varphi\Longleftrightarrow\varphi$ 在 $\mathfrak{A}$ 的每一逻辑可能世界上

真. (6)$\mathfrak{A}\models\Box^c\varphi\Longleftrightarrow\varphi$ 在 $\mathfrak{A}$ 的每一因果可能世界上真.

上述定义的模型被 Burks 称为模态模型. 称因果陈述句逻辑的一公式 φ 是有效式（记为 $\models\varphi$），若它在每一模态模型中真.

定理 7.1 若 φ 是因果陈述句逻辑的定理，则 $\models\varphi$.

该定理的证明如通常一样，只须证明因果陈述句逻辑的公理是有效的，推理规则保持有效性，而这是显然的.

我们称一逻辑系统 S 是一致的，若不存在一公式 φ 使得 φ 和 $\neg\varphi$ 都是 S 的定理.

定理 7.2 因果陈述句逻辑是一致的.

证明：令 φ 是因果陈述句逻辑的定理，据定理 7.1，$\models\varphi$，因此 $\not\models\neg\varphi$，据定理 7.1，$\neg\varphi$ 不是因果陈述句逻辑的定理.

Burks 并没有证明完全性定理，由于 Burks 的逻辑与模态逻辑的最小正规系统内容相差无几，因此可用类似模态逻辑典范模型的方法来证明完全性定理. 考虑因果条件句逻辑的非定理 φ，显然根据通常的方式 $\neg\varphi$ 能扩张成极大一致集 Φ. 这样的极大一致集可以看作是一个逻辑可能世界. 我们定义

$$\varphi\in\Phi\Longleftrightarrow\varphi\text{ 在 }\Phi\text{ 上真.}$$

易得 $\not\models\varphi$，因此完全性定理成立. 这里不象模态逻辑那样，还要在极大一致集之间定义通达关系，所以证明异常简单. 不过这一方法自然地扩大了 Burks 的逻辑可能世界的概念. Burks 建立在基本合取之上的逻辑可能世界的概念显得过于狭窄，因为它对语言的限制较多. Burks 的这一思想来自 Carnap，但 Carnap 之所以要如此建立状态描述（他的可能世界）是为了在状态描述中分配先验概率，而 Burks 的逻辑在这个阶段仍是模态逻辑，并不涉及概率值的分配问题，所以用极大一致集作为逻辑可能世界更好.

2. 具体解释

Burks 认为用因果陈述句逻辑的公式可以模型化（model）包含因果陈述句的演绎论证，这种用因果陈述句逻辑来模型化自然论证（日常生活中的论证或自然科学的论证）的方法可以具体解释因果陈述句逻辑。在此 Burks 讨论了如何用因果陈述句逻辑的有效论证来模型化因果律、因果虚拟句、因果倾向句和原因-结果陈述句（后三者依赖因果律）。

首先 Burks 定义了两种新的蕴涵关系：非悖论因果蕴涵（npc）和省略的因果蕴涵（ec）。

(1) $\varphi \mathrm{npc} \psi \triangleq \Box^{c}(\varphi \to \psi) \wedge \Diamond(\varphi \wedge \neg\psi) \wedge \Diamond(\neg\varphi \wedge \psi) \wedge \Diamond^{c}\varphi \wedge \neg\Box^{c}\psi$.

(2) $X(x) \mathrm{ec}\, Y(x) \triangleq (\exists Z)(Z(x) \wedge (Z(x) \wedge X(x)\, \mathrm{npc}\, Y(x)))$，（规定 npc 和 ec 的结合力小于∧，下同）。

(3) $\varphi\, \mathrm{ec} \psi \triangleq (\exists\varphi)(\varphi \wedge (\varphi \wedge \psi \mathrm{npc} \theta))$.

注意：在 §1 我们给出的形成规则中不允许对谓词变元和陈述句变元进行量化，而这里给出的(2)和(3)则允许，所以定义(2)和(3)中的谓词量词和陈述句量词是因果陈述句逻辑的非正式扩张。因此，类似于个体量词的情况，我们可以在因果陈述句逻辑中增加下列存在概括规则和存在例证(instantiation)规则：

规则 7.4（存在概括规则）

(1) $Z(x) \wedge (Z(x) \wedge X(x) \mathrm{npc} Y(x)) / X(x) \mathrm{ec} Y(x)$.

(2) $\varphi \wedge (\varphi \wedge \psi \mathrm{npc} \theta) / \psi \mathrm{ec} \theta$.

规则 7.5（存在例证规则）

(1) $X(x) \mathrm{ec} Y(x) / X(x) \to Y(x)$.

(2) $\varphi\, \mathrm{ec} \psi / \varphi \to \psi$.

易证，我们有如下定理成立。

定理 7.3

(1) $\vdash(\varphi \mathrm{npc} \psi) \rightarrow \neg \square(\varphi \rightarrow \psi)$.

(2) $\vdash(\varphi \mathrm{npc} \psi) \rightarrow \square^{c}(\varphi \rightarrow \psi)$.

(3) $\vdash(x)(\varphi \mathrm{npc} \psi) \rightarrow (x) \square^{c}(\varphi \rightarrow \psi)$.

(4) $\vdash \neg((\varphi \mathrm{npc} \psi) \wedge(\varphi \mathrm{npc} \neg \psi))$.

(5) $\vdash \neg(\varphi \mathrm{npc} \psi \wedge \neg \psi)$.

定理 7.4

(1) $(\varphi \mathrm{ec} \psi) \rightarrow(\varphi \rightarrow \psi)$.

(2) $(\varphi \wedge \psi \mathrm{ec} \theta) \rightarrow(\varphi \rightarrow(\psi \mathrm{ec} \theta))$.

(3) $(\varphi \wedge \psi \mathrm{npc} \theta) \rightarrow(\varphi \rightarrow(\psi \mathrm{ec} \theta))$.

(4) $(\psi \mathrm{ec} \theta) \wedge(\psi \mathrm{ec} \neg \theta) \rightarrow(\exists \varphi)(\varphi \wedge \neg \Diamond^{c}(\varphi \wedge \psi))$.

(5) $((X(x) \mathrm{ec} Y(x)) \wedge(X(x) \mathrm{ec} \neg Y(x))$

$\rightarrow(\exists Z)(Z(x) \wedge \neg \Diamond^{c}(Z(x) \wedge Y(x)))$.

(6) $\neg(\varphi \mathrm{ec} \psi \wedge \neg \psi)$.

注意：定理 7.4 中的各结论前面不用"$\vdash$"是因为它们是因果陈述句逻辑的非正式扩张。上述定理刻划了因果陈述句的许多重要的特性。

我们还可以通过下表对五种蕴涵号做一比较。

蕴涵的种类		当"若…，则…"被给定的蕴涵号置换时，下列表达式是定理吗?		
种类	符号	(若 $\varphi\wedge\psi$，则 θ)→(φ→(若 ψ，则 θ))	¬((若 ψ，则 θ)∧(若 ψ，则 ¬θ))	¬(若 ψ，则 $\theta\wedge\neg\theta$)
实质蕴涵	$\rightarrow$	是	不是	不是
逻辑蕴涵	$\prec$	不是	不是	不是
因果蕴涵	$\overset{c}{\rightarrow}$	不是	不是	不是
非悖论因果蕴涵	npc	不是	是	是
省略的因果蕴涵	ec	是	不是	是

Burks 认为自然规律分两类：一类是概然的，可以用下一节的归纳概率演算来模型化，一类涉及因果必然性，所以称

为因果律．用非悖论因果蕴涵来模型化涉及因果律的论证是非常恰当的．例如，考虑下列论证，其中大前提表达了一个因果律：

所有的黄金(G)物都在王水(A)中溶解(D)，

这一戒指(r)不是在王水中的黄金物，

所以，若这一戒指是在王水中的黄金物，则它将溶解．

表述"$G(r)\wedge A(r)$"的小前提在现实世界是假的．结论用的是虚拟句，这是为了表示结论是关于"$G(r)\wedge A(r)$"为真的非现实的因果可能世界的陈述句．我们用已建立起来的概念把上述论证符号化：

$(x)(G(x)\wedge A(x)\text{npc}D(x))$，

$\neg(G(r)\wedge A(r))$，

$\therefore\ \neg(G(r)\wedge A(r))\wedge(G(r)\wedge A(r)\text{npc}D(r))$．

可以证明这个论证在因果陈述句逻辑中是有效的．比较下列无效论证，其中大前提表达了现今世界的一个偶然事实：

在这个书桌(D)上的所有书(B)都是意大利文(I)的，

这(t)不是那个书桌上的书，

所以，若它是那个书桌上的书，则它是意大利文的．

这个论证可以符号化为：

$(x)(B(x)\wedge D(x)\rightarrow I(x))$，

$\neg(B(t)\wedge D(t))$，

$\therefore\ \neg(B(t)\wedge D(t))\wedge(B(t)\wedge D(t)\text{npc}I(t))$．

这个形式化论证不是有效的，因为实质全称句"$(x)(B(x)\wedge D(x)\rightarrow I(x))$"完全是关于现实世界的，并没有谈及任意可能的非现实世界．通过构造下列抽象模型易证这个论证是无效的：令现实世界是 $B(t)\wedge\neg D(t)\wedge\neg I(t)$，另一个因果可能世界是 $B(t)\wedge D(t)\wedge\neg I(t)$．

现在我们再考察一个用省略的因果蕴涵来模型化因果倾

向句的例子．假设某人的戒指是金的，因为金子在王水中是可溶解的，所以下列句子是真的

某人的戒指在王水中是可溶解的． (7.2)

Burks 称类似 (7.2) 的句子为因果倾向句．类似的因果倾向句还有"a 是易燃的"，"b 是可伸展的"，等等．因果倾向句在广义上还包括能力、潜能等因素．

陈述句(7.2)与下列虚拟句是同义的，或近似同义的．

某人的戒指是这样的：若它被放入王水，则它会溶解． (7.3)

虽然这句话为真的根据是戒指是金的，但句子中没有明确提到金子的事，若这个戒指是银的，这句话也为真．这就提示我们应该用下列句子来模型化上述(7.2)和(7.3)：

某人的戒指(r)具有一种属性 X，使得

$$X(r)\wedge A(r)\mathrm{npc}D(r). \quad (7.4)$$

用存在谓词量词来量化，就是

$$(\exists X)(X(r)\wedge(X(r)\wedge A(r)\mathrm{npc}D(r))),$$

即
$$A(r)\mathrm{ec}D(r).$$

作为说明，现在我们从前提"黄金在王水中溶解"和"某人的戒指是金的"推导出因果倾向句"某人的戒指在王水中是可溶解的"．

(1) $(x)(G(x)\wedge A(x)\mathrm{npc}D(x))$，前提；

(2) $G(r)$，前提；

(3) $G(r)\wedge A(r)\mathrm{npc}D(r)$，据(1)和谓词公理 7.1；

(4) $G(r)\wedge(G(r)\wedge A(r)\mathrm{npc}D(r))$，据(2)，(3)；

(5) $(\exists X)(X(r)\wedge(X(r)\wedge A(r)\mathrm{npc}D(r)))$，据(4)和规则 7.4；

(6) $A(r)\mathrm{ec}D(r)$，据(5)和 ec 的定义．

显然上述论证在扩张的非正式的因果陈述句逻辑中是有效

的.

倾向公式"$(\exists X)(X(r)\wedge(X(r)\wedge A(r)\mathrm{npc}D(r)))$"的一个重要特性是它陈述了所有这样的可能性. 如果人们知道这个戒指在王水中是可溶解的，那么他知道这个戒指有某种属性 X 以满足因果条件 $X(r)\wedge A(r)\mathrm{npc}D(r)$，但他不需要知道这种属性是什么. 因此因果倾向句的本质属性就是，它断定存在一个属性而没有明确地辨别这一属性. Burks 认为一个好的模型必须保持这种特性，所以"$G(r)$"和"$G(r)\wedge(x)(G(x)\wedge A(x)\mathrm{npc}D(x))$"都不是因果倾向句"某人的戒指在王水中可溶解"的好模型.

总之，Burks 认为对因果律、因果虚拟句、因果倾向句和原因一结果关系的陈述句用非悖论因果蕴涵句或省略的因果蕴涵句来模型化是恰当的，由此表述的论证通常是因果陈述句逻辑（或它的非正式扩张）中的有效论证，这就是 Burks 对因果陈述句逻辑的具体解释.

§3 因果必然性与归纳概率

1. 时空系统与归纳法

Burks 认为重复性在归纳法中起重要作用（如枚举归纳法），因此它既预设可重复因素，又预设重复性能发生其中的时空系统. 时空系统有两种：一种是自然时空系统（例如，一个容器中的现实气体（不是理想气体）），一种是人工定义的系统（例如，胞腔自动体（cellular automata））. 胞腔自动体的空间可以分隔为有穷个胞腔，它的时间维度可以分成持续的瞬间. 胞腔自动体的时空结构是连续的自然时空系统的一个离散模型.

Burks 把属性区别为非指标属性（如"黄"，"时间"，"男

人")和指标属性(如"现在","这里","这本书","那种情况"),因此表示这些属性的符号也分为非指标符号和指标符号. 时空系统中的基本个体是点或区域,在这些点或区域上,一元非指标属性可以被例证.为了简化,我们假设一胞腔自动体有有穷个一元非指标的基本属性. 一个胞腔状态由一个合取式来描述,相对每一基本属性,这个合取式陈述了它是否存在(类似基本合取). 自动体中的每一胞腔 c 的邻域 $N(c)$ 由 c 和其他与 c 有公共边的胞腔组成. 我们假设一胞腔在时间 $t+1$ 时的状态只依赖它在 t 时相邻的胞腔的状态, 因此胞腔自动体由建立在胞腔空间中的齐一的邻域上的"转移规则"(连续律)来支配.

概率胞腔自动体(或因果胞腔自动体)由"概率(或因果)转移规则"来支配,这个规则对赋于邻域 $N(c)$的每一胞腔状态,给出胞腔 c 的可能的后继状态上的概率分布(或给出 c 的一个或多个后继状态). 一个因果胞腔自动体称为决定论的,若每一时间状态(即对每一胞腔赋于一胞腔状态的赋值)有一唯一的后继,否则称为非决定论的. 实质上,这些因果自动体是因果陈述句逻辑的模态模型.每一个基本个体就是一个胞腔-瞬间,每一胞腔状态的析取是一胞腔属性(类似 Carnap 的 Q 性质). 逻辑可能世界由对所有胞腔-瞬间指派胞腔状态的赋值组成, 而因果可能世界是满足因果转移规则和初始状态规则(它规定一个因果可能世界的初始时间状态)的逻辑可能世界.因果胞腔自动体将用于构造因果陈述句逻辑的因果模型.

Burks 认为时空相邻性在归纳法中起重要的作用. 在科学中考虑的自然规律绝大多数是相邻性的(例外的有牛顿的万有引力律),所以这里假设自然规律都是相邻的. 相邻的因果律可以在因果胞腔自动体中模型化, 因为这些自动体的转移规则是建立在邻域关系 $N(c)$ 之上的因果规则. 为了在因

果陈述句逻辑中把这些规则符号化，令 $P\{N(c),t\}$ 是当时间 t 时关于胞腔 $N(c)$ 的诸原子陈述句的任意一致但非重言式的真值函数，即$P\{N(c),t\}$在 t 时把属性 P 指派给区域 $N(c)$. 同样，$Q\{c,t'\}$在时间 t' 把属性 Q 指派给胞腔 c，则

$$(c)(t)(P\{N(c),t\}\text{npc}Q(c,t')) \qquad (7.5)$$

表达了一个相邻的因果律(其中 (c),(t) 是全称量词)，它描述了在一个时间的间隔因果影响直接从一个胞腔传到它的邻域. 决定论的或非决定论的胞腔自动体的转移规则等价形如(7 5) 的相邻因果律的一个一致集. 注意，这些规律具有自然规律的唯一性、模态性和齐一性.

除了相邻性，自然规律还有一个重要的空间特性：局部性或整体性. 若一个规律作用的范围是相对小的空间，则称它为局部的. 若它作用的范围是无限大的空间，则称为是整体的. 相邻的因果律是局部的. 在时空上齐一的局部规律能推出整体规律. 相邻作用大于遥远作用的相邻局部规律称为准局部规律.

由于观察和实验都是局部活动，归纳探究和证实之所以成功，因为它预设自然律是准局部且在归纳上是简单的. 这个预设最易用胞腔模型来解释. 假设一个科学家能用一个有大量胞腔状态的决定论的胞腔自动体来观察和实验. 然而他不知道该自动体的转移规则，打算从观察证据中把它推导出来. 他最有可能是遇到这种情况，该自动体的转移规则由许多相邻的因果律组成，其中每一个在下列意义上是简单的：该因果律把邻域 $N(c)$ 的一个一般的属性与胞腔 c 的一个一般属性相连. 因此运用改变因果关系相关性质的方法就能确定这一规律系统. 这样的因果律就认为“归纳上是简单的”.

在从局部证据到整体结论的归纳推理中，有三个预设起重要作用，一个是因果齐一性原则：若非指标属性之间的因

果联系在一个时空域上成立，则它在整个时空域上成立．第二个是因果存在原则：某个时空域受准局部规律支配，至少对其大部是如此．第三个是有限变元原则：自然规律依赖一个有穷的非指标一元属性集，其中每一属性要求在极小时空区域中被例证．

2．Burks 的归纳逻辑理论

Burks 认为归纳推理就是从前提 d 到结论 c 的概然推理，把它符号化就成为 “$P(c,d)=x$”，其中“c”，“d”是陈述句变元，“d”的取值范围限制在逻辑可能的陈述句上，即“d”不能是矛盾句或逻辑上不可能的陈述句，“x”是数值变元，取[0,1]闭区间中的实数为值．Burks 把形如“$P(c,d)=x$”的表达式称为原子归纳概率陈述句．Burks 的条件归纳概率演算由下列公理组成：

公理 7.5 若$\Box(e\rightarrow \neg(c\wedge d))$，则
$$P(c\vee d,e)=P(c,e)+P(d,e).$$

公理 7.6 $P(c\wedge d)=P(c,e)\cdot P(d,c\wedge e)$.

公理 7.7 若$\Box(d\rightarrow c)$，则 $P(c,d)=1$.

公理 7.8 若$\Box(e\longleftrightarrow f)$，则 $P(c,e)=P(c,f)$.

若引入以下定义：

若 d 是逻辑真的，则 $P(c)\triangleq P(c,d)$，

则上述系统就成为无条件归纳概率演算．

Burks 认为归纳逻辑就是判定原子归纳概率陈述句真值的规则系统．显然，除了一些极限情况，上述演算不能确定大部分原子归纳陈述句的真值状况（所以上述公理系统应该说是一个公设系统），因此这些归纳概率演算只是完备的归纳逻辑的一部分．对于 Carnap 的语言系统 $\mathscr{L}_N^\pi$，Burks 在此讨论了三种归纳逻辑：(1)标准归纳逻辑：即 Carnap 的 c^*系统再加上类比规则，和改变因果相干性质的方法．(2)随机归

纳逻辑：即赋于状态描述的概率相等.(3)逆向归纳逻辑：对$\mathscr{L}_N^{\pi}$中每一状态描述 SD, $P(\mathrm{SD}) = I(\mathrm{SD}) \Big/ \sum_{i=1}^{2^{\pi N}} I(\mathrm{SD}_i)$，其中 $I(\mathrm{SD})$ 表示与 SD 同构的状态描述的数目．对逆向归纳逻辑来说，正证据的增加反而使归纳概率减少.对随机归纳逻辑来说，正证据的增加不影响归纳概率。对标准归纳逻辑来说，正证据的增加使归纳概率增加，因此实际工作的科学家主要使用标准归纳逻辑.

3. 因果必然性和归纳概率的联系

因为因果必然性是自然律、因果律、因果虚拟句、因果倾向句等的本质特性，所以对归纳法的完备解释必须说明因果陈述句如何被证实，归纳概率在这个证实过程中究竟起什么作用。Burks 认为，由于实际使用的归纳推理的规则构成标准归纳逻辑，所以上述问题就变成：因果必然陈述句如何在标准归纳逻辑中证实.

我们知道第二小节中使用的形式模型 $\mathfrak{A}_{\mathscr{L}}$（就是 Carnap 的语言系统 $\mathscr{L}_N^{\pi}$)是用以研究归纳逻辑的，它包含一个逻辑可能世界集(状态描述集)，通过把无条件概率(相当 Carnap 的正则测度)赋于逻辑可能世界来建立标准归纳逻辑等模型.这类模型可以称为归纳法的统计模型，它们解释了归纳法中重复性的作用，但它们不包括因果模态词，所以不能解释因果必然陈述句的证实，因为 $\mathfrak{A}_{\mathscr{L}}$ 不包含因果可能世界集，所以因果必然性概念在其中是不能表达的． 而在本章 §2 的抽象解释中引入的模态模型包含了因果可能世界集，因此它可以表达因果必然性概念，但不能表达归纳概率的概念．此外这两类模型都没有时空系统，而我们看到的因果律在归纳法上的某些重要特性(例如齐一性和相邻性)预设时空框架作用.

因此 Burks 要建立一种新的模型来解释因果律，这种模

型应该把标准归纳逻辑的统计模型的可重复特性、模态模型的模态性和胞腔自动体的时空组织结合起来. Burks 把这种模型称为标准归纳逻辑的因果模型. 它可分为两个部分：模态结构以及对此结构赋值归纳概率.

模态结构由一个有穷的“可能因果系统”的集合组成，其中每一可能因果系统包含一个(一般是无穷的)因果可能世界集. 一个标准归纳逻辑的因果模型的模态结构可由下列图表给出.

	可能因果系统 所有的因果胞腔自动体具有 m 个胞腔状态 且满足因果齐一性和因果存在预设				
	$\mathscr{S}_a$ (现实的) ($\square^c$)	$\mathscr{S}_1$ ($\square_1^c$)	$\mathscr{S}_2$ ($\square_2^c$)	… …	$\mathscr{S}_M$ ($\square_M^c$)
因果可能世界 这些时间状态的进程 满足该自动体的转 移规则	u_{a1} (现实的) u_{a2} u_{a3} ⋮	u_{11} u_{12} u_{13} ⋮	u_{21} u_{22} u_{23} ⋮	…	u_{M1} u_{M2} u_{M3} ⋮

在定义一个模型的可能因果系统之前，我们先用一种一般的方式来描述它们. 这些可能因果系统就是因果胞腔自动体，它们分别表示自然界和其他可能的世界. 它们都有时空框架,满足三个归纳预设：因果齐一性、因果存在和有限变元预设. 因为空间通常是三维的，且时间在过去和未来两个方向是无穷的，所以我们考虑的胞腔自动体有三个空间维度和一个在两个方向无穷的时间维度. 因此空间分成立体的胞腔，时间分成离散的瞬间 ($t=\cdots,-2,-1,0,1,2,\cdots$). 一个给定的模型的所有因果自动体都有相同数目的 m 个胞腔状

态．这些自动体只是相对于其转移规则而不同．因为转移规则等价于一个有穷的相邻的因果律的集合，所以它们只是相对于支配它们的因果律而不同．

现在我们把一个模型的可能因果系统定义为那些满足因果齐一性预设、因果存在预设和有限变元原则，并且有m个胞腔状态的胞腔自动体．由第三个预设可知胞腔状态的数目为有穷．我们可以用相同的术语把这里的因果模型和前面的形式模型、模态模型联系起来．前面的模型由个体常元和Q谓词构成的Q句组成，由此构成了逻辑可能世界和因果可能世界．在因果模型中，胞腔-瞬间就是个体常元指称的对象，胞腔状态就是用Q句来描述．逻辑可能世界就是时间状态的两个方向无穷的进程，其中时间状态就是胞腔状态对胞腔的赋值．注意一个给定模型的每一可能因果系统包含相同的逻辑可能世界集，因为该模型的每一胞腔自动体依赖相同的有穷胞腔状态集．最后，一个可能因果系统的因果可能世界由满足该系统的转移规则的逻辑可能世界构成．自然界是由所有的事物(事件或实体)构成的单一时空系统，其中的事物有时空位置并且受基本规律支配．在标准归纳逻辑的因果模型中(参见上页表)，自然界由现实因果系统$\mathscr{S}_a$表示，其他的可能因果系统$\mathscr{S}_1, \cdots, \mathscr{S}_M$是对应于自然界在归纳上可能的系统，因此它们满足归纳法的三个预设，从而在实际的归纳过程中予以考虑．$\mathscr{S}_a$中的一个世界表示我们现实的世界，因此称为这个模型的现实世界，记为u_{a1}．$\mathscr{S}_a, \mathscr{S}_1, \cdots, \mathscr{S}_M$中的每一个系统都是因果陈述句逻辑的模型，所以都定义了一个因果必然算子．我们仍用"$\Box^c$"表示$\mathscr{S}_a$中的因果必然，用新的符号"$\Box_1^c$"，…，"$\Box_M^c$"分别表示$\mathscr{S}_1, \cdots, \mathscr{S}_M$中的因果必然．因为"$\Box^c$"是用现实因果系统定义，所以下列等价成立：其中$\varphi$是非模态陈述句，

$\Box^c\varphi$ 真 $\Longleftrightarrow$ φ 在 $\mathscr{S}_a$ 的每一逻辑可能世界中真. (7.6)

φ 真 $\Longleftrightarrow$ φ在现实世界 u_{a1} 中真. (7.7)

这就完成了对标准归纳逻辑因果模型第一部分的描述. 注意,因果模型在下面几个方面是有穷的,这对归纳法是关键所在: 该模型的胞腔自动体的胞腔状态的数目m是固定的. 一个胞腔 c 的邻域 $N(c)$ 是有穷的, 它只能有有穷个不同的状态. 对 $N(c)$ 的每一状态, 转移规则给出 c 的一个或多个后继状态. 因为只涉及有穷多个状态, 所以转移规则原则上可以用一有穷表来表达. 可能的传移规则的数目和可能的因果系统的数目都是有穷的. 但另一方面, 一个可能因果系统可以包含无穷多个因果可能世界.

标准归纳逻辑的因果模型的第二部分由无条件归纳概率的赋值构成, 这个赋值分两步: 对该模型的可能因果系统的赋值以及对这些系统中的因果可能世界的赋值. 这个赋值根据无差别原则: 所有的可能因果系统赋予同等的归纳概率, 在每一系统中所有的因果可能世界也赋予同等的归纳概率. 因为一共有 $M+1$ 个可能因果系统,所以每个可能因果系统被赋于 $1/(M+1)$ 的归纳概率. 因为在每一可能因果系统中一般有无穷多个因果可能世界, 所以不能直接把相同的概率赋于给它们, 而是赋于出现在某个有穷时空区域中的诸可能世界的部分. 由于这样的赋值与因果律的证实过程相连, 所以我们在下面的过程中讨论.

因为标准归纳逻辑的因果模型是从因果胞腔自动体中构造起来的, 所以该因果模型的基本规律就是下列形式的相邻因果律:

$(c)(t)(P\{N(c),t\}\mathrm{npc}Q(c,t'))$, 其中$t'=t+1$. (7.8)

形如 (7.8)的陈述句表达了可能因果律,它在某些可能因果系统中真, 在另一些系统中假. 一个可能的因果系统由转移规

则定义，对 $N(c)$ 的每一状态，该规则给出 c 的一个或多个后继状态，所以它可以确定(7.8)是否真. 因为 $N(c)$ 的每一状态能表达为属性 P, c 的每一状态能表达为属性 Q，因此一个可能因果系统的转移规则等价一个有穷的可能因果律集. 在现实因果系统中成立的因果律表示自然律，称为现实因果律. 科学家用观察和实验来确定哪些可能的规律是现实的，即哪个可能因果系统是现实因果系统. 但是直接观察只限于现实世界，而规律定义的是因果可能世界集，所以 Burks 还要解决关于现实世界的观察如何确证关于因果可能世界的陈述句.

(7.8)中的因果全称句断定了，在现实因果系统的每一世界中有某种联系成立，因此它演绎蕴涵相应的实质全称句

$$(c)(t)(P\{N(c),t\} \rightarrow Q(c,t')). \qquad (7.9)$$

(7.9)演绎蕴涵特例陈述句：对时空框架的每一胞腔-瞬间 $\langle c_i,t_i \rangle$

$$P\{N(c_i),t_i\} \rightarrow Q(c_i,t_i'). \qquad (7.10)$$

通过在 t_i 时对 $N(c_i)$ 和在 t_i' 时对 c_i 的直接观察能确证或反驳(7.10). 直接反驳(7.8)是简单的：若(7.10)假则(7.8)假. 确证(7.8)的情况就更复杂. 从直观上显然(7.10)的成立在某种程度上确证了(7.8)，用概率论的术语说就是(7.8)相对(7.10)的后验概率大于(7.8)的先验概率.

我们来寻求导致上述结果的因果模型可能世界的概率. 令 $\mathfrak{A}$ 是一因果模型，我们用下列方式把 $\mathfrak{A}$ 归约为有穷模型 $\mathfrak{A}_f$. 考虑某个有穷时空区域 R 使得 R 大到足以包括人们以往搜集的观察证据和在可以预见的未来人们能搜集到的所有观察证据. 用 $\mathfrak{A}$ 的在 R 中的每一因果可能世界的部分作为可能世界，就构成有穷模型 $\mathfrak{A}_f$ 的因果可能世界. 虽然区域 R 相当大，但它是有穷的，它只包含有穷多个胞腔，而每个胞腔只有有穷

多个胞腔状态，所以 $\mathfrak{A}_f$ 只包含有穷多个因果可能世界。因此容易对每一可能世界等量地赋于一非零的无条件归纳概率（参见[16]，pp. 629—630）。这就说明了特例(7.10)如何在有穷因果模型中确证(7.8)。通过以下两点根据 Burks 推测，上述确证过程能扩张到无穷时空。(1) 归纳概率赋值表达了在不确定条件下，在某些方面人们活动的倾向。这些活动表达了人们的选择或偏好。人的绝大多数兴趣在范围上是有穷的，充分大的有穷时空区域包含这些兴趣的所有对象。归纳法依赖人的有限的兴趣，因此在解释因果必然陈述句的证实时使用有穷时空区域是令人满意的。但这个解释不包括与长驻性相关的归纳证实。(2) 对因果可能系统赋于概率存在一个极限过程，这在直观上是合乎情理的。这个极限过程解释了一个非模态特例如何确证一个模态规律。归纳法的统计模型说明了对这样的赋值存在极限过程。这些事实说明上述特例确证的解释可以扩张到无穷时空，也说明对确证过程做适当的概括能运用于因果模型。

Burks 的模态归纳逻辑十分丰富也十分繁杂。Burks 想给归纳法和归纳逻辑一个坚实的基础，然而他的工作并没有比 Carnap 的工作前进多少，他引入时空系统，胞腔自动体等一系列概念，除了使事情复杂化，并没有带来多少新成果，在这些方面似乎还不如 Cohen。

第八章 概率语义学

在《哲学逻辑手册》第一卷中(参见[17]),H. Leblanc 提出了概率语义学. 虽然在我们看来他的工作只是为一阶逻辑的弱化形式(参见下面的语言和公理 8.4)建立一种区别于标准语义学(Tarski 语义学)的概率语义学,但是这项工作毕竟是在 80 年代初做的,因此有许多新的有意义的成果,值得我们了解和讨论,至少能使我们更清楚地认识到概率的本性是语义的.从而建立一个形式概率语义学(对概率概念除了数学刻划不加别的说明)不仅是完全可能的,而且是非常有用的.这也反过来使我们了解归纳逻辑的使命以及它与概率逻辑的区别.

§1 一元概率语义学

语言 $\mathscr{L}$ 的初始符号是

(1) 可数个 n($n \geqslant 1$)元谓词: $Q_1, Q_2, \cdots$.

(2) 可数个个体常元,假设它们以字母表顺序排列: $t_1, t_2, \cdots$.

(3) 可数个个体变元: $x_1, x_2, \cdots$.

注意 $\mathscr{L}$ 没有函数符号. 逻辑符号、形成规则和联结词定义如通常,其中 $\neg, \wedge$ 是初始联结词. 我们用"φ","ψ","θ"表示公式,"S"表示公式集. 此外还用 $\bigwedge\limits_{i=1}^{n} \varphi_i$ 缩写"$((\cdots(\varphi_1 \wedge \varphi_2) \wedge \cdots) \wedge \varphi_n)$". 我们称 $\varphi(t_1/x), \varphi(t_2/x), \cdots$ 是$\mathscr{L}$

的量化句 $(x)\varphi$ 在 $\mathscr{L}$ 中的代入特例。

$\mathscr{L}$ 的公理是所有具有下列公理 8.1—8.6 形式的陈述句：

公理 8.1 $\varphi\rightarrow(\varphi\wedge\varphi)$.

公理 8.2 $(\varphi\wedge\psi)\rightarrow\varphi$.

公理 8.3 $(\varphi\rightarrow\psi)\rightarrow(\neg(\psi\wedge\theta)\rightarrow\neg(\theta\wedge\varphi))$.

公理 8.4 $\varphi\rightarrow(x)\varphi$，其中 x 在 φ 中不出现.

公理 8.5 $(x)\varphi\rightarrow\varphi(t/x)$，其中 t 是 $\mathscr{L}$ 的个体常元.

公理 8.6 $(x)(\varphi\rightarrow\psi)\rightarrow((x)\varphi\rightarrow(x)\psi)$.

$\mathscr{L}$ 的规则如下：

规则 8.1 $\varphi\rightarrow\psi,\varphi/\psi$，其中 $\varphi\rightarrow\psi,\varphi$ 是陈述句.

规则 8.2 若 φ 是 $\mathscr{L}$ 的公理，则 $(x)\varphi(x/t)$ 也是公理，其中 t 是 $\mathscr{L}$ 的个体常元.

我们称 $\mathscr{L}^+$ 是 $\mathscr{L}$ 的一个(个体)常元扩张，若 $\mathscr{L}^+$ 是任意和 $\mathscr{L}$ 完全相同的语言，除了在 $\mathscr{L}$ 的个体常元之外还增加可数个个体常元。当然，根据该定义，$\mathscr{L}$ 也是 $\mathscr{L}$ 的常元扩张. 这里假设 $\mathscr{L}^+$ 的常元用某种字母表顺序排列，用 "t_i^+" $(i=1,2,\cdots)$ 来指称.

Leblanc 继承了 Keynes 的观点，认为概率就是合理相信度，是陈述句的特性，而不是陈述的形式。在本节，我们考虑一元概率函数 P^+，它取 $\mathscr{L}^+$ 的单个陈述句为变目，函数值是闭区间 $[0,1]$ 中的实数。我们用下列 7 个条件来刻划 $\mathscr{L}^+$ 的任意概率函数 P^+，使得一个合理认知主体可以以程度分别为 $P^+(\varphi_1),P^+(\varphi_2),\cdots$ 同时相信 $\mathscr{L}^+$ 的陈述句 $\varphi_1,\varphi_2,\cdots$，其中 φ_1 是按字母表顺序排列的第一个陈述句，φ_2 是第二个，…．

条件 8.1 $0\leqslant P^+(\varphi)$.

条件 8.2 $P^+(\neg(\varphi\wedge\neg\varphi))=1$.

条件 8.3 $P^+(\varphi)=P^+(\varphi\wedge\psi)+P^+(\varphi\wedge\neg\psi)$.

条件 8.4 $P^+(\varphi)\leqslant P^+(\varphi\wedge\varphi)$.

条件 8.5 $P^+(\varphi\wedge\phi)\leqslant P^+(\phi\wedge\varphi)$.

条件 8.6 $P^+(\varphi\wedge(\phi\wedge\theta))\leqslant P^+((\varphi\wedge\phi)\wedge\theta)$.

条件 8.7 $P^+(\varphi\wedge(x)\phi)=\lim\limits_{j\to\infty}P^+\left(\varphi\wedge\bigwedge\limits_{i=1}^{j}\phi(t_i^+/x)\right)$.

条件 8.1—8.7 可以看作是概率逻辑的真值表（参见第二章§3)，也可把 P^+看作是对一阶句子的一种解释(模型).

现在我们引入下列重要定义.

定义8.1 令 ST($\mathscr{L}$)是 $\mathscr{L}$ 的陈述句集，称 $\varphi\in$ ST($\mathscr{L}$) 在概率意义上逻辑真$\Longleftrightarrow$对 $\mathscr{L}$ 的任意常元扩张 $\mathscr{L}^+$和 $\mathscr{L}^+$的任意概率函数 P^+，$P^+(\varphi)=1$. 令 $S\subseteq$ST($\mathscr{L}$)，称在概率意义上 S 逻辑推出 $\varphi\Longleftrightarrow$若对任意 $\mathscr{L}^+$和 $\mathscr{L}^+$的任意 P^+，对每一 $\phi\in B$，$P^+(\phi)=1$，则 $P^+(\varphi)=1$.

应该注意，在下列定理中，若把“P”换作“P^+”，定理仍成立，从而使后面的定理 8.43 和 8.48 也成立.

定理 8.1 $P(\varphi\wedge\phi)\leqslant P(\varphi)$.

证明：据条件 8.1 和 8.3.

定理 8.2 $P(\varphi)=P(\varphi\wedge\varphi)$.

据条件 8.4 和定理 8.1，上述定理显然成立.

定理 8.3 $P(\varphi\wedge\phi)=P(\phi\wedge\varphi)$.

据条件 8.5，上述定理显然成立.

定理 8.4 $P((\varphi\wedge\phi)\wedge\theta)=P(\varphi\wedge(\phi\wedge\theta))$.

证明：据条件 8.6，只须证

$$P((\varphi\wedge\phi)\wedge\theta)\leqslant P(\varphi\wedge(\phi\wedge\theta)),$$

而 $P((\varphi\wedge\phi)\wedge\theta)$

$=P(\theta\wedge(\varphi\wedge\phi))$ （据定理 8.3）

$\leqslant P((\theta\wedge\varphi)\wedge\phi)$ （据条件 8.6）

$\leqslant P(\phi\wedge(\theta\wedge\varphi))$ （据定理 8.3）

$\leqslant P((\phi\wedge\theta)\wedge\varphi)$ （据条件 8.6）

$\leqslant P(\varphi \wedge (\psi \wedge \theta))$ （据定理 8.3）.

定理 8.5 $P(\varphi) = P(\psi \wedge \varphi) + P(\neg\psi \wedge \varphi)$.

据条件 8.3 和定理 8.3，上述定理显然成立.

定理 8.6 若 $P(\varphi) \leqslant 0$，则 $P(\varphi) = 0$.

据条件 8.1，上述定理显然成立.

定理 8.7 $P(\varphi \wedge \psi) \leqslant P(\psi)$.

据定理 8.1 和 8.3，上述定理成立.

定理 8.8 $P(\varphi \wedge \neg\varphi) = 0$.

据条件 8.3 和定理 8.2，上述定理成立.

定理 8.9 $P((\varphi \wedge \neg\varphi) \wedge \psi) = 0$.

据定理 8.1，8.6，8.8，上述定理成立.

定理 8.10 $P(\varphi) = P(\neg(\psi \wedge \neg\psi) \wedge \varphi)$.

据定理 8.5 和 8.9，上述定理成立.

定理 8.11 $P(\neg\varphi) = 1 - P(\varphi)$.

证明： $P(\neg\varphi)$

$= P(\neg(\varphi \wedge \neg\varphi) \wedge \neg\varphi)$ （据定理 8.10）

$= P(\neg(\varphi \wedge \neg\varphi)) - P(\neg(\varphi \wedge \neg\varphi) \wedge \varphi)$ （据条件 8.3）

$= 1 - P(\neg(\varphi \wedge \neg\varphi) \wedge \varphi)$ （据条件 8.2）

$= 1 - P(\varphi)$ （据定理 8.10）.

据条件 8.1 和定理 8.11，我们有

定理 8.12 $P(\varphi) \leqslant 1$.

另外，还可推出

定理 8.13 若 $P(\varphi) \geqslant 1$，则 $P(\varphi) = 1$.

定理 8.14 $P(\neg\varphi \wedge (\varphi \wedge \psi)) = 0$.

证明： $P(\neg\varphi \wedge (\varphi \wedge \psi))$

$= P((\neg\varphi \wedge \varphi) \wedge \psi)$ （据定理 8.4）

$\leqslant P(\neg\varphi \wedge \varphi)$ （据定理 8.1）

$= 0$ （据定理 8.8 和 8.6）.

从定理 8.5 和 8.14，推出

定理 8.15 $P(\varphi\wedge\psi)=P(\varphi\wedge(\varphi\wedge\psi))$.

从定理 8.11 推出

定理 8.16 $P(\varphi\rightarrow\psi)=1\Longleftrightarrow P(\varphi\wedge\neg\psi)=0$.

从定理 8.16 和条件 8.3，我们有

定理 8.17 $P(\varphi\rightarrow\psi)=1\Longleftrightarrow P(\varphi)=P(\varphi\wedge\psi)$.

定理 8.18 $P(\varphi\rightarrow(\varphi\wedge\varphi))=1$.

从定理 8.2 和 8.15，推出

$P(\varphi)=P(\varphi\wedge\varphi)=P(\varphi\wedge(\varphi\wedge\varphi))$.

据定理 8.17,就有 $P(\varphi\rightarrow(\varphi\wedge\varphi))=1$，此定理与公理 8.1 对应。

定理 8.19 $P(\varphi\wedge\psi\rightarrow\varphi)=1$.

证明： 据定理 8.17,8.15 和 8.3,定理成立,此定理与公理 8.2 对应。

定理 8.20 $P(\varphi\wedge\psi\rightarrow\psi)=1$.

证明：据定理 8.17,8.3,8.4,8.9 以及条件 8.3,我们有

$$
\begin{aligned}
P(\varphi\wedge\psi)&=P((\varphi\wedge\psi)\wedge\psi)+P((\varphi\wedge\psi)\wedge\neg\psi)\\
&=P((\varphi\wedge\psi)\wedge\psi)+P(\varphi\wedge(\psi\wedge\neg\psi))\\
&=P((\varphi\wedge\psi)\wedge\psi)+P((\psi\wedge\neg\psi)\wedge\varphi)\\
&=P((\varphi\wedge\psi)\wedge\psi).
\end{aligned}
$$

定理 8.21 若 $P(\varphi\rightarrow\psi)=1$ 且 $P(\psi\rightarrow\theta)=1$，则 $P(\varphi\rightarrow\theta)=1$.

证明：假设 $P(\varphi\rightarrow\psi)=1$，根据定理 8.4,8.7 和 8.16，则

$$
\begin{aligned}
0&=P(\varphi\wedge\neg\psi)\\
&\geqslant P((\neg\theta\wedge\varphi)\wedge\neg\psi).
\end{aligned}
$$

再假设 $P(\psi\rightarrow\theta)=1$,则据定理 8.1,8.3,8.4,8.6,8.16,我们有

$$\begin{aligned}0 &= P(\psi \wedge \neg\theta) \\ &\geqslant P((\psi \wedge \neg\theta) \wedge \varphi) \\ &= P((\neg\theta \wedge \varphi) \wedge \psi).\end{aligned}$$

因此，据条件 8.3，$P(\neg\theta \wedge \varphi) = 0$，据定理 8.3，$P(\varphi \wedge \neg\theta) = 0$，因此据定理 8.16，得证。

从定理 8.17，8.4 和 8.1 推出

定理 8.22 若 $P(\varphi \to (\psi \wedge \theta)) = 1$，则 $P(\varphi \to \psi) = 1$。

定理 8.23 $P(\varphi \wedge \neg\neg\psi) = P(\varphi \wedge \psi)$。

证明：据条件 8.3，有

$$P(\varphi \wedge \neg\psi) + P(\varphi \wedge \neg\neg\psi) = P(\varphi \wedge \neg\psi) + P(\varphi \wedge \psi).$$

因此定理成立。

定理 8.24 若 $P(\varphi \to (\psi \to \theta)) = 1$ 且 $P(\varphi \to \psi) = 1$，则 $P(\varphi \to \theta) = 1$。

证明：先假设 $P(\varphi \to (\psi \to \theta)) = 1$，从定理 8.3，8.4，8.16 和 8.23，则

$$\begin{aligned}0 &= P(\varphi \wedge (\psi \wedge \neg\theta)) \\ &= P((\neg\theta \wedge \varphi) \wedge \psi).\end{aligned}$$

假设 $P(\varphi \to \psi) = 1$，从定理 8.4，8.6，8.7 和 8.16 推出

$$\begin{aligned}0 &\geqslant P(\neg\theta \wedge (\varphi \wedge \neg\psi)) \\ &= P((\neg\theta \wedge \varphi) \wedge \neg\psi).\end{aligned}$$

因此据条件 8.3 和定理 8.3，$P(\varphi \wedge \neg\theta) = 0$，据定理 8.16，$P(\varphi \to \theta) = 1$。

定理 8.25 若 $P(\varphi \to \psi) = 1$ 且 $P(\varphi \to \theta) = 1$，则 $P(\varphi \to (\psi \wedge \theta)) = 1$。

证明：假设 $P(\varphi \to \psi) = 1$，据定理 8.17，定理 8.3 和 8.4 以及条件 8.3有

$$P(\varphi) = P(\varphi \wedge \psi)$$

$$= P(\varphi\wedge(\psi\wedge\theta)) + P(\neg\theta\wedge(\varphi\wedge\psi)),$$

再假设 $P(\varphi\rightarrow\theta)=1$，则根据定理 8.1，8.3，8.4，8.6，8.16，有

$$\begin{aligned} 0 &= P(\neg\theta\wedge\varphi) \\ &\geqslant P((\neg\theta\wedge\varphi)\wedge\psi) \\ &= P(\neg\theta\wedge(\varphi\wedge\psi)). \end{aligned}$$

因此，$P(\varphi)=P(\varphi\wedge(\psi\wedge\theta))$，据定理 8.17，$P(\varphi\rightarrow(\psi\wedge\theta))=1$。

定理 8.26 若 $P(\varphi\rightarrow(\psi\wedge\theta))=1$，则 $P(\varphi\rightarrow\theta)=1$。

证明：假设 $P(\varphi\rightarrow(\psi\wedge\theta))=1$，据定理 8.17 和 8.3 等，$P(\varphi)\leqslant P(\theta\wedge\varphi)$，因此据定理 8.22 的证明中同样的推理方法，推出 $P(\varphi\rightarrow\theta)=1$。

定理 8.27 若 $P(\varphi\rightarrow\psi)=1$ 且 $P(\varphi\rightarrow\neg\psi)=1$，则 $P(\varphi)=0$。

证明：假设前提成立，则据定理 8.16，

$P(\varphi\wedge\neg\psi)=0, P(\varphi\wedge\neg\neg\psi)=0,$

因此据条件 8.3，$P(\varphi)=0$。

定理 8.28 若 $P((\varphi\wedge\psi)\rightarrow\theta)=1$，则 $P(\varphi\rightarrow(\psi\rightarrow\theta))=1$。

证明：假设前提成立，则据定理 8.16，$0=P((\varphi\wedge\psi)\wedge\neg\theta)$，据定理 8.4 和 8.23，$0=P(\varphi\wedge\neg\neg(\psi\wedge\neg\theta))$，据定理 8.17，$P(\varphi\rightarrow(\psi\rightarrow\theta))=1$。

定理 8.29 $P((\varphi\rightarrow\psi)\rightarrow(\neg(\psi\wedge\theta)\rightarrow\neg(\theta\wedge\varphi))=1$（=公理 8.3）。

证明：(1) 据定理 8.20 和 8.19，有

$P((((\varphi\rightarrow\psi)\wedge\neg(\psi\wedge\theta))\wedge(\theta\wedge\varphi))\rightarrow(\theta\wedge\varphi))=1,$

$P((\theta\wedge\varphi)\rightarrow\theta)=1.$

因此据定理 8.21，

$$P((((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) \to \theta) = 1.$$

(2) 同理用定理 8.20 代替 8.19，有

$$P((((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) \to \varphi) = 1, \quad (8.1)$$

而据 8.19 和 8.22，

$$P((((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) \to (\varphi \to \psi)) = 1.$$

因此据定理 8.24 和(8.1)，

$$P((((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) \to \psi) = 1.$$

(3) 据(1)，(2)和定理 8.25，我们有

$$P((((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) \to (\psi \land \theta)) = 1, \quad (8.2)$$

而据定理 8.19 和 8.26，

$$P((((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) \to \neg(\psi \land \theta)) = 1.$$

因此据定理 8.27 和(8.2)，

$$P(((\varphi \to \psi) \land \neg(\psi \land \theta)) \land (\theta \land \varphi)) = 0.$$

从定理 8.23 和 8.16 推出

$$P(((\varphi \to \psi) \land \neg(\psi \land \theta)) \to \neg(\theta \land \varphi)) = 1.$$

由定理 8.28，我们有

$$P((\varphi \to \psi) \to (\neg(\psi \land \theta) \to \neg(\theta \land \varphi))) = 1.$$

定理 8.30 若 $P(\varphi) = 1$ 且 $P(\varphi \to \psi) = 1$，则

$$P(\psi) = 1.$$

证明：假设前提成立，则据定理 8.17，$P(\varphi \land \psi) = 1$，据定理 8.7 和 8.13，$P(\psi) = 1$，此定理对应于规则 8.1.

定理 8.31 若 $\vdash_0 \varphi$（即 φ 的证明只用公理 8.1—8.3 和规则 8.1)，则 $P(\varphi) = 1$.

证明据定理 8.18，8.19，8.29，8.30 且施归纳于 φ 的证明的长度.

定理 8.32 若$\vdash_0 \varphi \longleftrightarrow \psi$，则 $P(\varphi) = P(\psi)$.

证明：若$\vdash_{\circ} \varphi \longleftrightarrow \psi$，则据定理 8.31，

$$P(\varphi \rightarrow \psi) = 1 \text{ 且 } P(\psi \rightarrow \varphi) = 1.$$

因此据定理 8.17，

$$P(\varphi) = P(\varphi \wedge \psi), \ P(\psi) = P(\psi \wedge \varphi).$$

定理 8.33 $P((x)\varphi) = \lim_{n \to \infty} P\left(\bigwedge_{i=1}^{j} \varphi(t_i/x)\right)$.

证明：据条件 8.7，$P(\neg(\varphi \wedge \neg\varphi) \wedge (x)\varphi) = \lim_{j \to \infty} P(\neg(\varphi \wedge \neg\varphi) \wedge \bigwedge_{i=1}^{j} \varphi(t_i/x))$，据定理 8.10 和极限定义，就有定理 8.33。

定理 8.34 若 $P(\varphi) = 1$ 且 $P(\psi) = 1$，则

$$P(\varphi \wedge \psi) = 1.$$

证明：假设 $P(\varphi) = 1$，则 $P(\varphi \wedge \psi) + P(\varphi \wedge \neg\psi) = 1$。假设 $P(\psi) = 1$，则 $P(\neg\psi) = 0$，据定理 8.7 和 8.6，$P(\varphi \wedge \neg\psi) = 0$，因此据定理 8.34，$P(\varphi \wedge \psi) = 1$。

定理 8.35 若对 $i = 1, \cdots, j$，$P(\varphi_i) = 1$，则

$$P\left(\bigwedge_{i=1}^{j} \varphi_i\right) = 1.$$

定理 8.36 若对从 1 开始的每一 i，$P(\varphi(t_i/x)) = 1$，则 $P((x)\varphi) = 1$。

证明：据定理 8.35 和 8.33 以及极限定义，定理成立。

定理 8.37 若对从 1 开始的每一 i，$P(\varphi \rightarrow \psi(t_i/x)) = 1$，则 $P(\varphi \rightarrow (x)\psi) = 1$。

证明：假设前提成立，则据定理 8.25，且对 j 作归纳，

$$P(\varphi \rightarrow \bigwedge_{i=1}^{j} \psi(t_i/x)) = 1.$$

因此，根据定理 8.17，条件 8.7，有

$$P(\varphi)=P(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x))$$

$$=\lim_{j\to\infty}P(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x))$$

$$=P(\varphi\wedge(x)\psi).$$

因此据定理 8.17，$P(\varphi\to(x)\psi)=1$。

定理 8.38 $P(\varphi\to(x)\varphi)=1$。

证明：因为 x 在 φ 中不出现，所以 $\varphi(t_i/x)$ 就是 φ，因此据定理 8.32 和极限定义，

$$\lim_{j\to\infty}P(\varphi\wedge\bigwedge_{i=1}^{j}\varphi(t_i/x))=P(\varphi),$$

据条件 8.7 和定理 8.17，$P(\varphi\to(x)\varphi)=1$，这定理与公理 8.4 对应。

定理 8.39 $P((x)\varphi\to\varphi(t/x))=1$.

证明：令 t 是 $\mathscr{L}$ 的第 k 个个体，对 $j\geqslant k$，$\vdash_0(\varphi(t/x)\wedge\bigwedge_{i=1}^{j}\varphi(t_i/x))\longleftrightarrow\bigwedge_{i=1}^{j}\varphi(t_i/x)$，因此据定理 8.32 和极限定义，我们有

$$\lim_{j\to\infty}P(\varphi(t/x)\wedge\bigwedge_{i=1}^{j}\varphi(t_i/x))=\lim_{j\to\infty}P\left(\bigwedge_{i=1}^{j}\varphi(t_i/x)\right).$$

据条件 8.7 和定理 8.33，有

$$P(\varphi(t/x)\wedge(x)\varphi)=P((x)\varphi),$$

据定理 8.3 和 8.17，

$$P((x)\varphi\to\varphi(t/x))=1.$$

此定理对应于公理 8.5

定理 8.40 $P((x)(\varphi\to\psi)\to((x)\varphi\to(x)\psi))=1$.

证明：令 t 是 $\mathscr{L}$ 的任意常元，据定理 8.19，

$$P(((x)(\varphi\to\psi)\wedge(x)\varphi)\to(x)(\varphi\to\psi))=1.$$

而据定理 8.39，

$$P((x)(\varphi \to \psi) \to (\varphi \to \psi)(t/x)) = 1.$$

因此据定理 8.21,

$$P(((x)(\varphi \to \psi) \wedge (x)\varphi) \to (\varphi \to \psi)(t/x)) = 1.$$

同理用定理 8.20 代替 8.19, 有

$$P(((x)(\varphi \to \psi) \wedge (x)\varphi) \to \varphi(t/x)) = 1.$$

因此据定理 8.24,

$$P(((x)(\varphi \to \psi) \wedge (x)\varphi) \to \psi(t/x)) = 1.$$

从定理 8.37 和对 t 的假设, 推出

$$P(((x)(\varphi \to \psi) \wedge (x)\varphi) \to (x)\psi) = 1.$$

据定理 8.28,定理成立,此定理对应于公理 8.6.

定理 8.41 若 φ 是 $\mathscr{L}$ 的公理,则 $P(\varphi) = 1$.

证明: 假设 φ 是 $\mathscr{L}$ 的公理, 反复使用规则 8.2, φ 成为

$$(x_1)(x_2)\cdots(x_n)(\psi(x_1, x_2, \cdots, x_n/t_1, t_2, \cdots, t_n)),$$

其中 $n \geqslant 0$, 且 ψ 是公理 8.1—8.6 中之一. 我们对 n 运用归纳法来证明 $P(\varphi) = 1$.

当 $n = 0$ 时,则 $\varphi(=\psi)$ 是公理 8.1—8.6 中之一, 因此据定理 8.18—8.19,8.29 和 8.38—8.40, $P(\varphi) = 1$.

当 $n > 0$ 时,显然对从 1 开始的每一 i,

$$((x_2)\cdots(x_n)(\psi(x_1, x_2, \cdots, x_n/t_1, t_2, \cdots, t_n)))(t_i/x_1)$$

构成 $\mathscr{L}$ 的公理. 根据归纳假设,对从 1 开始的每一 i,

$$P(((x_2)\cdots(x_n)(\psi(x_1, x_2, \cdots, x_n/t_1, t_2, \cdots, t_n)))(t_i/x_1)) = 1.$$

因此据定理 8.37, $P((x_1)(x_2)\cdots(x_n)(\psi(x_1, x_2, \cdots, x_n/t_1, t_2, \cdots, t_n))) = 1$,所以 $P(\varphi) = 1$,定理证毕.

定理 8.42 若 $S \vdash \varphi$, 且对所有 $\psi \in S$, $P(\psi) = 1$, 则 $P(\varphi) = 1$.

证明: 对 φ 运用归纳法证明. 假设在 $\mathscr{L}$ 中从 S 到 φ 的证明是 $\psi_1, \psi_2, \cdots, \psi_p(=\varphi)$,且假设对所有 $\psi \in S, P(\psi) =$

1. 若 $\phi_i(1 \leqslant i \leqslant p) \in S$,据假设, $P(\phi_i) = 1$. 若 ϕ_i 是 $\mathscr{L}$ 的公理,据定理 8.41, $P(\phi_i) = 1$. 若 ϕ_i 是从这个证明更前的部分根据规则 8.2 得到的,则据我们熟悉的证明论事实,定理 8.36 和归纳假设, $P(\phi_i) = 1$. 若 ϕ_i 是据规则 8.1 得到,则据定理 8.30. 因此 $P(\phi_p) = 1$,即

$$P(\varphi) = 1.$$

因为在上述定理中,用"P^+"置换"P"定理仍成立,所以我们有以下重要定理:

定理 8.43 (概率语义学中的强可靠定理)若 $S \vdash \varphi$, 则在概率意义上, S 逻辑推出 φ.

要证明定理 8.43 的逆定理, 还需以下定义和定理.

定义 8.2 令 $S \subseteq \mathrm{ST}(\mathscr{L})$, 称函数 P_S 是 $\mathscr{L}$ 中 S 的相伴概率,若 P_S 满足下列条件: 对任意 $\varphi \in \mathrm{ST}(\mathscr{L})$, 若 $S \vdash \varphi$, 则 $P_S(\varphi) = 1$, 否则 $P_S(\varphi) = 0$.

定义 8.3 称 S 在 $\mathscr{L}$ 中(不)一致,若不存在(存在)$\mathscr{L}$中一陈述句 φ 使得 $S \vdash \varphi \wedge \neg\varphi$. 称 S 在 $\mathscr{L}$ 中极大一致, 若 S 在 $\mathscr{L}$ 中一致,且对 $\mathscr{L}$ 的任一不在 S 中的陈述句 φ,$S \cup \{\varphi\}$ 在$\mathscr{L}$ 中不一致. 称 S 在 $\mathscr{L}$ 中 ω 完全, 若对 $\mathscr{L}$ 的任意量化句 $(x)\varphi$,对 $\mathscr{L}$ 的每一常元 t, $S \vdash \varphi(t/x)$,则 $S \vdash (x)\varphi$.

定理 8.44 令 $S \subseteq \mathrm{ST}(\mathscr{L})$, P_S 是 $\mathscr{L}$ 中 S 的相伴概率,则

(1) P_S 满足条件 8.1—8.2,8.4—8.6.

(2) 若 S 在 $\mathscr{L}$ 中极大一致, 则 P_S 满足条件 8.1—8.6.

(3) 若 S 在 $\mathscr{L}$ 中 ω 完全,则 P_S 满足条件 8.1—8.2,8.4—8.7.

(4) 若 S 在 $\mathscr{L}$ 中极大一致且 ω 完全, 则 P_S 构成 $\mathscr{L}$ 的概率函数(即 P_S 满足条件 8.1—8.7).

证明: (1)可以直接从 P_S 的定义和 $\mathscr{L}$ 中关于证明论的

基本事实，其细节留给读者补充。

(2) 假设 S 在 $\mathscr{L}$ 中极大一致，且先假设 $P_S(\varphi)=1$，则（据 P_S 的定义）$S\vdash\varphi$。因为 $S\vdash\varphi\to(\varphi\wedge\psi)\vee(\varphi\wedge\neg\psi)$，所以$(\varphi\wedge\psi)\vee(\varphi\wedge\neg\psi)\in S$，因此 $\varphi\wedge\psi\in S$ 或 $\varphi\wedge\neg\psi\in S$，从而或 $S\vdash\varphi\wedge\psi$ 或 $S\vdash\varphi\wedge\neg\psi$ 成立。因此 $P_S(\varphi\wedge\psi)$ 和 $P_S(\varphi\wedge\neg\psi)$ 中有一个为 1，另一个为 0，亦即 $P_S(\varphi)=P_S(\varphi\wedge\psi)+P_S(\varphi\wedge\neg\psi)$，再假设 $P_S(\varphi)=0$，则 $S\nvdash\varphi$，因此 $S\nvdash\varphi\wedge\psi$ 且 $S\nvdash\varphi\wedge\neg\psi$，$P_S(\varphi\wedge\psi)=P_S(\varphi\wedge\neg\psi)=0$，$P_S(\varphi)=P_S(\varphi\wedge\psi)+P_S(\varphi\wedge\neg\psi)$。

(3) 假设 S 在 $\mathscr{L}$ 中 ω 完全，$P_S(\varphi\wedge(x)\psi)=1$，则 $S\vdash\varphi\wedge(x)\psi$。因此，对从 1 开始的每一 j，$S\vdash\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)$，即得，$P_S\left(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)\right)=1$，亦即 $\lim_{j\to\infty}P_S\left(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)\right)=1$。再假设 $P_S(\varphi\wedge(x)\psi)=0$，则 $S\nvdash\varphi\wedge(x)\psi$，因此或 $S\nvdash\varphi$ 或 $S\nvdash(x)\psi$（或两者皆不成立）。若 $S\nvdash\varphi$，则对任意 j，$S\nvdash\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)$，因此对任意 j，$P_S\left(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)\right)=0$。另一方面，若 $S\nvdash(x)\psi$，则至少对 $\mathscr{L}$ 的一个常元 t，$S\nvdash\psi(t/x)$，这是因为 S 在 $\mathscr{L}$ 中 ω 完全。因此，存在一 k（k 大于等于 t 的脚码）使得对从 k 开始的每一 j，$S\nvdash\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)$，所以存在一 k 使得对从 k 开始的每一 j，$P_S\left(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)\right)=0$，因此 $\lim_{j\to\infty}P_S\left(\varphi\wedge\bigwedge_{i=1}^{j}\psi(t_i/x)\right)=0$，亦即 P_S 满足条件 8.7。

定理 8.45 令 $S \subseteq \mathrm{ST}(\mathscr{L})$ 且在 $\mathscr{L}$ 中极大一致且 ω 完全，P_S 是 S 在 $\mathscr{L}$ 中的相伴概率，$\varphi \in \mathrm{ST}(\mathscr{L})$，则 $\varphi \in S \Longleftrightarrow P_S(\varphi) = 1$.

证明： 若 $\varphi \in S$，则 $S \vdash \varphi$，因此据 P_S 的定义，$P_S(\varphi) = 1$. 若 $\varphi \notin S$，则据 S 的极大一致性，$\neg\varphi \in S$，因此

$$P_S(\neg\varphi) = 1.$$

据定理 8.44(4)和定理 8.11，因此 $P_S(\varphi) = 0$，即

$$P_S(\varphi) \neq 1.$$

定义 8.4 称一个常元外在于 $\varphi \in \mathrm{ST}(\mathscr{L})$，若它不在 φ 中出现. 称一个常元外在于 $S \subseteq \mathrm{ST}(\mathscr{L})$，若它对 S 中每一元素都是外在的. 称 $S \subseteq \mathrm{ST}(\mathscr{L})$ 在 $\mathscr{L}$ 中无穷可扩张，若 $\mathscr{L}$ 中有 ω 个常元外在于 S.

当 $S \subseteq \mathrm{ST}(\mathscr{L})$ 在 $\mathscr{L}$ 中一致且无穷可扩张，则存在一种方法（Henkin 方法）把 S 扩张到 $\mathrm{H}(S) \subseteq \mathrm{ST}(\mathscr{L})$，且使得 $\mathrm{H}(S)$ 在 $\mathscr{L}$ 中极大一致且 ω 完全. 这样的集合 $\mathrm{H}(S)$ 称为 S 在 $\mathscr{L}$ 中的 Henkin 扩张. 下面我们来简述构造 $\mathrm{H}(S)$ 的方法. 首先构造集合链：

$$S_0 \subseteq S_1 \subseteq \cdots \subseteq S_n \subseteq \cdots,$$

其中 $S_0 = S$，且对从 1 开始的每一 i，令 φ_n 是 $\mathscr{L}$ 的字母顺序排列的第 n 个陈述句. 若 $S_{n-1} \cup \{\varphi_n\}$ 不一致，则令 $S_n = S_{n-1}$. 若 $S \cup \{\varphi_n\}$ 在 $\mathscr{L}$ 中一致且 φ_n 不是 $\mathscr{L}$ 的否定的全称量化句 $\neg(x)\phi$，则令 $S_n = S_{n-1} \cup \{\varphi_n\}$. 若 $S_{n-1} \cup \{\varphi_n\}$ 在 $\mathscr{L}$ 中一致，且 φ_n 是 $\mathscr{L}$ 的否定的全称量化句 $\neg(x)\phi$，则令 $S_n = S_{n-1} \cup \{\varphi_n, \neg\phi(t/x)\}$，其中 t 是 $\mathscr{L}$ 中以字母顺序第一个外在于 $S_{n-1} \cup \{\varphi_n\}$ 的常元. 然后我们构造 $\mathrm{H}(S)$ 如下：

$$\mathrm{H}(S) = \bigcup_{i=0}^{\infty} S_i,$$

易证这样的 $\mathrm{H}(S)$ 是极大一致且 ω 完全的.

定理 8.46 令任意 $\varphi \in \mathrm{ST}(\mathscr{L})$，$S \subseteq \mathrm{ST}(\mathscr{L})$ 在 $\mathscr{L}$ 中无穷可扩张．若 $S \nvdash \varphi$，则存在 $\mathscr{L}$ 的一概率函数 P 使得对所有 $\phi \in S$，$P(\phi) = 1$ 但 $P(\varphi) \neq 1$.

证明：假设 $S \nvdash \varphi$，其中 S 在 $\mathscr{L}$ 中无穷可扩张，则 $S \cup \{\neg\varphi\}$ 在 $\mathscr{L}$ 中一致．据 Henkin 方法，$\mathrm{H}(S \cup \{\neg\varphi\})$ 在 $\mathscr{L}$ 中极大一致且 ω 完全，因此据定理 8.45，$\mathrm{H}(S \cup \{\neg\varphi\})$ 的每一元素被 $\mathscr{L}$ 中 $\mathrm{H}(S \cup \{\neg\varphi\})$ 的相伴概率 P 赋值为 1．而据定理 8.44(4)，这个相伴概率 P 构成 $\mathscr{L}$ 的概率函数，所以 P 指派 0 给 φ.

在 Henkin 方法和定理 8.11，8.44—8.45 中，用"$\mathscr{L}^+$"置换 $\mathscr{L}$，用"P^+"置换"P"，这些方法和定理仍成立，因此据定理 8.46，我们有

定理 8.47 S 和 φ 定义如定理 8.46，若对任意语言 $\mathscr{L}^+$ 和 $\mathscr{L}^+$ 的任意概率函数 P^+，若对所有 $\phi \in S, P^+(\phi) = 1$，则 $P^+(\varphi) = 1$，则 $S \vdash \varphi$.

若 S 在 $\mathscr{L}$ 中不是无穷可扩张的，则 $S \cup \{\neg\varphi\}$ 在 $\mathscr{L}^\infty$（$\mathscr{L}^\infty$ 是 $\mathscr{L}$ 的一个常元扩张，即 $\mathscr{L}^\infty$ 除了 $\mathscr{L}$ 的常元之外，还有 ω 个新常元）中仍是无穷可扩张的．因此若 $S \nvdash \varphi$，则据 Henkin 方法，定理 8.45，8.46 和 8.11，存在 $\mathscr{L}$ 的一个常元扩张 $\mathscr{L}^+(\mathscr{L}^\infty)$ 和 $\mathscr{L}^+$ 的一个概率函数 P^+（即 $\mathscr{L}^\infty$ 中 $\mathrm{H}(S \cup \{\neg\varphi\})$ 的相伴概率）使得对所有 $\phi \in S$，$P^+(\phi) = 1$ 但 $P^+(\varphi) \neq 1$．因此就有下列定理：

定理 8.48（概率语义学中 $\mathscr{L}$ 的强完全性定理）令任意 $S \subseteq \mathrm{ST}(\mathscr{L})$，任意 $\varphi \in \mathrm{ST}(\mathscr{L})$，若在概率意义上 S 逻辑推出 φ，则 $S \vdash \varphi$.

因此据定理 8.43 和令 $S = \varnothing$，又有

定理 8.49 令任意 $\varphi \in \mathrm{ST}(\mathscr{L})$，$\vdash \varphi \Leftrightarrow$ 在概率意义上 φ 逻辑真.

§2 真值函数与概率函数的关系

在本节我们讨论真值函数和一元概率函数的关系。$\mathscr{L}$ 的真值函数 $P:\mathrm{ST}(\mathscr{L})\to\{0,1\}$ 满足下列三个条件：

条件 8.8 $P(\neg\varphi)=1-P(\varphi)$.

条件 8.9 $P(\varphi\wedge\psi)=P(\varphi)\cdot P(\psi)$.

条件 8.10 $P((x)\varphi)=1\Longleftrightarrow$ 对 $\mathscr{L}$ 的每一常元 t, $P(\varphi(t/x))=1$.

显然，$\mathscr{L}$ 的真值函数是从 $\mathrm{ST}(\mathscr{L})$ 到满足条件 8.8—8.10 的实数的二值函数。反之，若 P 是一个从 $\mathrm{ST}(\mathscr{L})$ 到实数的，且满足条件 8.8—8.10 的二值函数，则 P 以 0 和 1 作为它的两个值，因此是 $\mathscr{L}$ 的真值函数。

定理 8.50 令 P 是从 $\mathrm{ST}(\mathscr{L})$ 到实数的二值函数，则有

(1) 若 P 满足条件 8.9，则 P 有 0 和 1 作为它的值。

(2) 若 P 满足条件 8.1—8.4（见§1），则 P 有 0 和 1 作为它的值。

证明：(1) 假设前提成立，则对从 1 开始的每一 n，$P\left(\bigwedge_{i=1}^{n}\varphi_i\right)=\prod_{i=1}^{n}P(\varphi_i)$，假设至少 P 的一个值是不等于 0 和 1 的某个实数 r，则对 $i\geqslant 3$，P 将有 $r,r^2,r^3,\cdots$ 这样的值，因此 P 不是二值函数，矛盾。

(2) 据条件 8.2，$P(\neg(\varphi\wedge\neg\varphi))=1$，据条件 8.1，8.3 和 8.4，$P(\varphi\wedge\neg\varphi)=0$（参见定理 8.8 的证明），因此条件 8.9 成立，所以：

定理 8.51 P 构成 $\mathscr{L}$ 的真值函数 $\Longleftrightarrow P$ 是从 $\mathrm{ST}(\mathscr{L})$ 到实数的满足条件 8.8—8.10 的二值函数。

因此条件 8.8—8.10 等价于标准的一阶语义学。

定理 8.52 令 P 是从 $ST(\mathscr{L})$ 到实数的二值函数，若 P 满足条件 8.8—8.10，则 P 满足条件 8.1—8.7.

证明：根据题设和定理 8.50(1)，P 只有 0 和 1 作为值．若 P 满足条件 8.8—8.10，则显然 P 满足条件 8.1—8.2，8.4—8.6．因为据条件 8.8—8.9，有

$$
\begin{aligned}
P(\varphi) &= (P(\varphi)\cdot P(\phi)) + P(\varphi) - (P(\varphi)\cdot P(\phi)) \\
&= P(\varphi\wedge\phi) + P(\varphi) - (P(\varphi)\cdot P(\phi)) \\
&= P(\varphi\wedge\phi) + P(\varphi)(1 - P(\phi)) \\
&= P(\varphi\wedge\phi) + P(\varphi)\cdot P(\neg\phi) \\
&= P(\varphi\wedge\phi) + P(\varphi\wedge\neg\phi),
\end{aligned}
$$

所以 P 满足条件 8.3.

假设 $P(\varphi\wedge(x)\phi) = 1$，据条件 8.9，$P(\varphi)\cdot P((x)\phi) = 1$，据定理 8.50(1)，$P(\varphi) = P((x)\phi) = 1$，因此据条件 8.10，对 $\mathscr{L}$ 中每一常元 t，$P(\phi(t/x)) = 1$．因此据条件 8.9，对从 1 开始的每一 j，$P\left(\varphi\wedge\bigwedge_{i=1}^{j}\phi(t_i/x)\right) = 1$，即 $\lim_{j\to\infty}P\left(\varphi\wedge\bigwedge_{i=1}^{j}\phi(t_i/x)\right) = 1$．再假设 $P(\varphi\wedge(x)\phi) = 0$，据条件 8.9，$P(\varphi)\cdot P((x)\phi) = 0$，因此或 $P(\varphi) = 0$ 或 $P((x)\phi) = 0$．若 $P(\varphi) = 0$，则据条件 8.9，对从 1 开始的每一 j，

$$P\left(\varphi\wedge\bigwedge_{i=1}^{j}\phi(t_i/x)\right) = 0,$$

因此，

$$\lim_{j\to\infty}P\left(\varphi\wedge\bigwedge_{i=1}^{j}\phi(t_i/x)\right) = 0.$$

若 $P((x)\phi) = 0$，则据条件 8.10，至少对 $\mathscr{L}$ 的一个常元 t，$P(\phi(t/x)) = 0$，因此据条件 8.9，存在一 k 使得从 k 开始的每一 j，

$$P\left(\varphi \wedge \bigwedge_{i=1}^{j} \psi\left(t_i/x\right)\right)=0,$$

因此

$$\lim_{j \to \infty} P\left(\varphi \wedge \bigwedge_{i=1}^{j} \psi\left(t_i/x\right)\right)=0,$$

亦即 P 满足条件 8.7。

要证明定理 8.52 的逆定理还需下列两定理。注意，在第一个定理中，P 可以是二值到 ω 值，但在第二个定理中，P 只能是二值的。

定理 8.53 令 P 是从 $\mathrm{ST}(\mathscr{L})$ 到实数的满足条件 8.1—8.5 的函数，则有

(1) $P(\neg\varphi)=1-P(\varphi)$ (=条件 8.8)。

(2) 若 $P(\varphi \wedge \psi)=1$，则 $P(\varphi)=P(\psi)=1$。

(3) 若 $P(\varphi \wedge \psi)=1$，则 $P(\varphi \wedge \psi)=P(\varphi) \cdot P(\psi)$。

证明：(1) 就是定理 8.11，而该定理的证明只用到条件 8.1—8.5。(2) 假设 $P(\varphi \wedge \psi)=1$，据定理 8.1 和 8.13，$P(\varphi)=1$，据定理 8.7 和 8.13，$P(\psi)=1$，而定理 8.1，8.13 和 8.7 的证明只用到条件 8.1—8.5。据(2)易证(3)成立。

定理 8.54 令 P 是从 $\mathscr{L}$ 的陈述句集到实数的满足条件 8.1—8.5 和 8.7 的二值函数，则有

(1))若 $P(\varphi \wedge \psi)=0$，则或 $P(\varphi)=0$ 或 $P(\psi)=0$。

(2) 若 $P(\varphi \wedge \psi) \neq 1$，则 $P(\varphi \wedge \psi)=P(\varphi) \cdot P(\psi)$。

(3) $P(\varphi \wedge \psi)=P(\varphi) \cdot P(\psi)$ (=条件 8.9)。

(4) $P((x)\varphi)=1 \Longleftrightarrow$ 对 $\mathscr{L}$ 的每一常元 t，$P(\varphi(t/x))=1$ (等价于条件 8.10)。

证明：(1) 假设 $P(\varphi \wedge \psi)=0$，据条件 8.3，$P(\varphi)=P(\varphi \wedge \neg\psi)$。再假设 $P(\varphi) \neq 0$，则据定理 8.50(2)，$P(\varphi)=1$，因此 $P(\varphi \wedge \neg\psi)=1$。由定理 8.7 和 8.13，$P(\neg\psi)=1$，

由定理 8.11，$P(\psi)=0$，因此或 $P(\varphi)=0$ 或 $P(\psi)=0$。(2)据定理 8.50(2)和本定理的(1)可证(2)成立。(3) 据定理 8.53(3)和本定理的(2)可推出(3)成立。(4)假设 $P((x)\varphi)=1$，据定理 8.33，$\lim_{j\to\infty} P\left(\bigwedge_{i=1}^{j} \varphi(t_i/x)\right)=1$。但由定理 8.50(2)，每一 $P\left(\bigwedge_{i=1}^{1} \varphi(t_i/x)\right)$，$P\left(\bigwedge_{i=1}^{2} \varphi(t_i/x)\right)$，…或等于 0 或等于 1。据定理 8.1，它们中每一个小于或等于前面一个，因此对 $j\geq 1$，$P\left(\bigwedge_{i=1}^{j} \varphi(t_i/x)\right)=1$。由定理 8.53(2)，对 $\mathscr{L}$ 的每一常元 t，$P(\varphi(t/x))=1$。再假设 $P((x)\varphi)\neq 1$，据定理 8.50(2)，$P((x)\varphi)=0$，由定理 8.33，

$$\lim_{j\to\infty} P\left(\bigwedge_{i=1}^{j} \varphi(t_i/x)\right)=0,$$

因此据定理8.50(2)，存在一 k 使得对 $j\geq k$，

$$P\left(\bigwedge_{i=1}^{j} \varphi(t_i/x)\right)=0.$$

因此从本定理 (1)，至少对 $\mathscr{L}$ 的一个常元 t，$P(\varphi(t/x))=0$，即 $P(\varphi(t/x))\neq 1$。

定理 8.55 令 P 是从 $\mathscr{L}$ 的陈述句集到实数的二值函数，若 P 满足条件 8.1—8.7，则 P 满足条件 8.8—8.10。

据定理 8.52 和 8.55，我们可得

定理 8.56 令 P 是从 $\mathrm{ST}(\mathscr{L})$ 到实数的二值函数，P 满足条件 8.8—8.10 $\Longleftrightarrow$ P 满足条件 8.1—8.7。

定理 8.57 P 构成 $\mathscr{L}$ 的真值函数 $\Longleftrightarrow$ P 构成 $\mathscr{L}$ 的二值概率函数。

可见，概率理论只是真值理论的一种概括 (Reichenbach 就持这种观点)，它允许真值函数(在满足条件 8.1—8.7 的意

义上理解）有从二值到 ω 值中的任意值．真值理论是概率理论的限制，它要求概率函数只有二值．值得注意：这里的概率函数不能取 ω_1 个(第一个不可数基数)值，但是二值概率函数(即真值函数)在一个基本的方面不同于多于二值的概率函数，即前者是外延的，而后者不是．我们这里称 $\mathscr{L}$ 的概率函数 P 是外延的，若 P 的值唯一依赖其变目中的陈述句的直接子陈述句的 P 的值．更形式地说，若 P 满足下列条件：

(i) 给定 $P(\varphi)=P(\varphi')$，则 $P(\neg\varphi)=P(\neg\varphi')$．

(ii) 给定 $P(\varphi)=P(\varphi')$ 和 $P(\psi)=P(\psi')$，则

$$P(\varphi\wedge\psi)=P(\varphi'\wedge\psi').$$

(iii) 对 $\mathscr{L}$ 的每一常元 t，若给定 $P(\varphi(t/x))=P(\varphi'(t/x))$，则 $P((x)\varphi)=P((x)\varphi')$．

因为据定理 8.11，$P(\neg\varphi)=1-P(\varphi)$，所以无论 P 有什么值，(i)都成立．当 P 是二值函数时，从定理 8.54(3)，$P(\varphi\wedge\psi)=P(\varphi)\cdot P(\psi)$，所以(ii)成立．又据定理 8.33，$P((x)\varphi)=\lim\limits_{j\to\infty}P\left(\bigwedge\limits_{i=1}^{j}\varphi(t_i/x)\right)$，所以(iii)成立．但是当 P 是多于二值时，(ii)从而(iii)不成立．例如，假设 $\varphi\in ST(\mathscr{L})$，$P(\varphi)=1/2$，因此 $P(\neg\varphi)=1/2$．据定理 8.2，$P(\varphi\wedge\varphi)=P(\varphi)=1/2$，而据定理 8.8，$P(\varphi\wedge\neg\varphi)=0$，所以不满足(ii)．因此在 $\mathscr{L}$ 的概率函数中，只有二值概率函数是外延的，其余的都不是．这两个结果是预料之中的，因为 P 在此被理解为合理相信的测度，当 P 是二值时，它只是真值函数，而真值函数是外延的．

§3 二元概率语义学

本节我们首先研究二元概率语义学，然后讨论它与一元

概率语义学的关系。

令 $\mathscr{L}^+$是 $\mathscr{L}$ 的任意常元扩张，称 P^+是 $\mathscr{L}^+$的二元概率函数，若 P^+是从 $\mathscr{L}^+$的每一陈述句对到实数的函数使得 P^+满足下列条件：

条件 8.11 存在 $\mathscr{L}^+$的一陈述句 φ 和一陈述句 ψ 使得 $P^+(\varphi,\psi)\neq 1$。

条件 8.12 $0\leqslant P^+(\varphi,\psi)$。

条件 8.13 $P^+(\varphi,\varphi)=1$。

条件 8.14 若存在 $\theta\in \mathrm{ST}(\mathscr{L}^+)$使得 $P^+(\theta,\psi)\neq 1$，则 $P^+(\neg\varphi,\ \psi)=1-P^+(\varphi,\psi)$。

条件 8.15 $P^+(\varphi\wedge\psi,\theta)=P^+(\varphi,\psi\wedge\theta)\cdot P^+(\psi,\theta)$。

条件 8.16 $P^+(\varphi\wedge\psi,\theta)=P^+(\psi\wedge\varphi,\theta)$。

条件 8.17 $P^+(\varphi,\psi\wedge\theta)=P^+(\varphi,\theta\wedge\psi)$。

条件 8.18 $P^+((x)\varphi,\psi)=\lim\limits_{j\to\infty}P^+\left(\bigwedge\limits_{i=1}^{j}\varphi(t_i^+/x),\psi\right)$。

从直观上解释，$P^+(\varphi_1,\psi)$，$P^+(\varphi_2,\psi)$，$\cdots$是一合理认知主体根据 $\psi\in \mathrm{ST}(\mathscr{L}^+)$同时相信 $\varphi_1,\varphi_2,\cdots\in \mathrm{ST}(\mathscr{L}^+)$的程度。

定义 8.5 对任意 $\mathscr{L}^+$的任意 P^+，称 $\psi\in \mathrm{ST}(\mathscr{L}^+)$是 P^+正规(P^+反常)的，若存在(不存在) $\varphi\in \mathrm{ST}(\mathscr{L}^+)$使得$P^+(\varphi,\psi)\neq 1$。

因此若ψ是 P^+反常的，则对任意 $\varphi\in \mathrm{ST}(\mathscr{L}^+)$，$P^+(\varphi,\psi)=1$。

定义 8.6 称在二元概率意义上 $\varphi\in \mathrm{ST}(\mathscr{L})$ 逻辑真，若对 $\mathscr{L}$ 的任意 $\mathscr{L}^+$ 和 $\mathscr{L}^+$ 的任意 P^+ 和任意 $\psi\in \mathrm{ST}(\mathscr{L})$，$P^+(\varphi,\psi)=1$。称在二元概率意义上 $S\subseteq \mathrm{ST}(\mathscr{L})$ 逻辑推出 φ，若对 $\mathscr{L}$ 的任意 $\mathscr{L}^+$，$\mathscr{L}^+$的任意 P^+和任意 $\psi\in \mathrm{ST}(\mathscr{L})$，如果对每一 $\theta\in S$，$P^+(\theta,\psi)=1$，那么 $P^+(\varphi,\psi)=1$。

注意：在下列定理中用“P^+”代换“P”，这些定理仍成立。

定理 8.58 $P(\varphi,\psi)\leqslant 1$。

证明：若ψ是P反常的，则据定义 8.5，$P(\varphi,\psi)=1$。假设ψ是P正规的，则据条件 8.12，$P(\neg\varphi,\psi)\geqslant 0$，因此据条件 8.14，$P(\varphi,\psi)\leqslant 1$。

定理 8.59 $P(\varphi,\psi\wedge\varphi)=1$。

证明：据条件 8.13，$P(\psi\wedge\varphi,\psi\wedge\varphi)=1$，从条件 8.15，$P(\psi,\varphi\wedge(\psi\wedge\varphi))\cdot P(\varphi,\psi\wedge\varphi)=1$，而据条件 8.12 和定理 8.58，$P(\psi,\varphi\wedge(\psi\wedge\varphi))\in[0,1]$ 且 $P(\varphi,\psi\wedge\varphi)\in[0,1]$，所以它们中每一个皆为 1。

定理 8.60 $P(\varphi,\varphi\wedge\psi)=1$。

证明据定理 8.59 和条件 8.17。

定理 8.61 令ψ是P正规的，则(1) $P(\neg\psi,\psi)=0$，(2) $P(\neg\psi\wedge\varphi,\psi)=0$。

证明：(1) 据条件 8.13，8.14 和关于ψ的假设，可证(1)成立。(2) 据条件 8.15，$P(\neg\psi\wedge\varphi,\psi)=P(\varphi\wedge\neg\psi,\psi)$，据条件 8.15，$P(\varphi\wedge\neg\psi,\psi)=P(\varphi,\neg\psi\wedge\psi)\cdot P(\neg\psi,\psi)$，据(1)即有，$P(\neg\psi\wedge\varphi,\psi)=0$。

定理 8.62 若不存在 $\neg\varphi\wedge\psi\in ST(\mathscr{L})$使得$P(\varphi,\neg\varphi\wedge\psi)=0$，则对每一这样的 ψ，$P(\varphi,\psi)=1$。

证明：假设存在 $\psi\in ST(\mathscr{L})$ 使得 $P(\varphi,\psi)\neq 1$，则ψ是P正规的，因此据定理8.61(2)，$P(\neg\psi\wedge\neg\varphi,\psi)=0$。据条件 8.15，$P(\neg\psi,\neg\varphi\wedge\psi)\cdot P(\neg\varphi,\psi)=0$，但因为 $P(\varphi,\psi)\neq 1$ 且ψ是P正规的，所以据条件 8.14，$P(\neg\varphi,\psi)\neq 0$，因此，$P(\neg\psi,\neg\varphi\wedge\psi)=0$，亦即 $\neg\varphi\wedge\psi$ 是P正规的。但据定理 8.60，$P(\neg\varphi,\neg\varphi\wedge\psi)=1$，从条件 8.14，$P(\varphi,\neg\varphi\wedge\psi)=0$。

定理 8.63 $P(\varphi\wedge\psi,\theta)\leqslant P(\psi,\theta)$.

证明：据条件 8.15，$P(\varphi\wedge\psi,\theta)=P(\varphi,\psi\wedge\theta)\cdot P(\psi,\theta)$,但由条件 8.12 和定理 8.58，$P(\varphi,\psi\wedge\theta)$ 和 $P(\psi,\theta)$ 都在区间[0,1]中，因此它们中间没有一个小于 $P(\varphi\wedge\psi,\theta)$，所以定理成立.

定理 8.64 $P(\varphi\wedge\psi,\theta)\leqslant P(\varphi,\theta)$.

定理 8.65 若 $P(\varphi\wedge\psi,\theta)=1$,则

$$P(\varphi,\theta)=P(\psi,\theta)=1.$$

证明：假设 $P(\varphi\wedge\psi,\theta)=1$,因此

$$P(\varphi\wedge\psi,\theta)=P(\psi\wedge\varphi,\theta)=P(\psi,\varphi\wedge\theta)\cdot P(\varphi,\theta),$$

据定理 8.64 和 8.58，$P(\varphi,\theta)=1$，又由定理 8.63 和 8.58，$P(\varphi,\theta)=1$.

定理 8.66 若 $P(\varphi\to\psi,\theta)=0$，则 $P(\varphi,\theta)=1$ 且 $P(\psi,\theta)=0$.

证明：若 $P(\varphi\to\psi,\theta)=0$,则$\theta$是$P$正规的，因此据"→"的定义和条件 8.14，$P(\varphi\wedge\neg\psi,\theta)=1$，又由定理 8.65，$P(\varphi,\theta)=P(\neg\psi,\theta)=1$. 再据条件 8.14，就有本定理成立.

定理 8.67 $P(\varphi\wedge\varphi,\psi)=P(\varphi,\psi)$.

据条件 8.15 和定理 8.60 可证定理成立.

定理 8.68 若θ是P正规的,则

$$P(\varphi,\theta)=P(\varphi\wedge\psi,\theta)+P(\varphi\wedge\neg\psi,\theta).$$

证明：若$\varphi\wedge\theta$是P正规的(否则,亦然)，据条件 8.14，

$$P(\psi,\varphi\wedge\theta)+P(\neg\psi,\varphi\wedge\theta)=P(\theta,\varphi\wedge\theta)+P(\neg\theta,\varphi\wedge\theta).$$

在上述等式两边乘上因子 $P(\varphi,\theta)$,则有

$$P(\psi,\varphi\wedge\theta)\cdot P(\varphi,\theta)+P(\neg\psi,\varphi\wedge\theta)\cdot P(\varphi,\theta)=P(\theta,\varphi\wedge\theta)\cdot P(\varphi,\theta)$$

$+ \hat{P}(\neg\theta, \varphi \wedge \theta) \cdot \hat{P}(\varphi, \theta).$

据条件 8.15，我们有 $P(\psi \wedge \varphi, \theta) + P(\neg\psi \wedge \varphi, \theta) = \hat{P}(\theta, \varphi \wedge \theta) \cdot P(\varphi, \theta) + P(\neg\theta \wedge \varphi, \theta)$，从定理 8.59，定理8.61(2)和假设，即可证得定理成立。

定理 8.69 若 $P(\varphi \to \psi, \theta) = 1$ 且 $P(\varphi, \theta) = 1$，则 $P(\psi, \theta) = 1$.

证明：若 θ 是 P 反常的，则据定义，$P(\psi, \theta) = 1$，因此假设 θ 是 P 正规的且 $P(\varphi \to \psi, \theta) = 1$，据条件 8.14 和"→"的定义，$P(\varphi \wedge \neg\psi, \theta) = 0$，据定理 8.68，$P(\varphi \wedge \psi, \theta) = P(\varphi, \theta)$，因为 $P(\varphi, \theta) = 1$，则 $P(\varphi \wedge \psi, \theta) = 1$，因此据定理 8.63，$P(\psi, \theta) = 1$，上述定理对应于规则 8.1.

定理 8.70 若 $P(\varphi, \theta) = P(\psi, \theta) = 1$，则 $P(\varphi \wedge \psi, \theta) = 1$.

证明：若 θ 是 P 反常的，则显然. 若 θ 是 P 正规的，且 $P(\varphi, \theta) = 1$，则据定理 8.68，$P(\varphi \wedge \psi, \theta) + P(\varphi \wedge \neg\psi, \theta) = 1$. 而据条件 8.14 和假设 $P(\psi, \theta) = 1$，$P(\neg\psi, \theta) = 0$，因此据定理 8.68 $P(\neg\psi \wedge \varphi, \theta) + P(\neg\psi \wedge \neg\varphi, \theta) = 0$，即 $P(\neg\psi \wedge \varphi, \theta) = 0$，所以据交换律和上述结果可得，$P(\varphi \wedge \psi, \theta) = 1$.

定理 8.71 若 φ 形如公理 8.1—8.6，则对每一 $\psi \in ST(\mathscr{L})$，$P(\varphi, \psi) = 1$.

证明：令前提成立，我们证明若存在 $\psi \in ST(\mathscr{L})$ 使得 $P(\varphi, \neg\varphi \wedge \psi) = 0$，则会产生矛盾. 因此不存在 $\psi \in ST(\mathscr{L})$ 使得 $P(\varphi, \neg\varphi \wedge \psi) = 0$，所以据定理 8.62，$P(\varphi, \psi) = 1$.

情况 1：φ 形如 $\varphi' \to \varphi' \wedge \varphi'$. 假设存在 $\psi \in ST(\mathscr{L})$ 使得 $P(\varphi, \neg\varphi \wedge \psi) = 0$，即 $P(\varphi' \to \varphi' \wedge \varphi', \neg\varphi \wedge \psi) = 0$. 则据定理 8.66，(1) $P(\varphi', \neg\varphi \wedge \psi) = 1$，(2) $P(\varphi' \wedge \varphi', \neg\varphi \wedge \psi) = 0$，但据(2)和定理 8.67 可知，$P(\varphi', \neg\varphi \wedge \psi) =$

$=0$,这与(1)矛盾.

情况 2：φ 形如 $\varphi'\wedge\psi'\rightarrow\varphi'$. 假设存在 $\psi\in\mathrm{ST}(\mathscr{L})$ 使得 $P(\varphi,\neg\varphi\wedge\psi)=0$. 据定理 8.66,我们有(1) $P(\varphi'\wedge\psi',\neg\varphi\wedge\psi)=1$. (2) $P(\varphi',\neg\varphi\wedge\psi)=0$. 由(1)和定理 8.65,得 $P(\varphi',\neg\varphi\wedge\psi)=1$,与(2)矛盾.

情况 3：φ 形如 $(\varphi'\rightarrow\psi')\rightarrow(\neg(\psi'\wedge\theta')\rightarrow\neg(\theta'\wedge\varphi'))$. 假设存在 $\psi\in\mathrm{ST}(\mathscr{L})$ 使得 $P(\varphi,\neg\varphi\wedge\psi)=0$. 据定理 8.66,(1) $P(\varphi'\rightarrow\psi',\ \neg\varphi\wedge\psi)=1$. (2) $P(\neg(\psi'\wedge\theta')\rightarrow\neg(\theta'\wedge\varphi'),\ \neg\varphi\wedge\psi)=0$. 但据(2)和定理 8.66, 又有,(3) $P(\neg(\psi'\wedge\theta'),\neg\varphi\wedge\psi)=1$. (4) $P(\neg(\theta'\wedge\varphi'),\neg\varphi\wedge\psi)=0$. 因此据条件 8.14, 我们又有,(5) $P(\theta'\wedge\varphi',\neg\varphi\wedge\psi)=1$. 而据定理 8.65,得(6) $P(\varphi',\neg\varphi\wedge\psi)=1$. 由(1)和定理 8.69,得 (7)$P(\psi',\neg\varphi\wedge\psi)=1$. 由条件 8.14 和(3), 有 (8) $P(\psi'\wedge\theta',\neg\varphi\wedge\psi)=0$. 因此据条件 8.16, (9) $P(\theta'\wedge\psi',\neg\varphi\wedge\psi)=0$. 而据(5)和定理 8.65, (10) $P(\theta',\ \neg\varphi\wedge\psi)=1$. 因此由(9), (10)和定理 8.68, 得 $P(\theta'\wedge\neg\psi',\ \neg\varphi\wedge\psi)=1$. 据定理 8.65, $P(\neg\psi',\ \neg\varphi\wedge\psi)=1$, 这与(7)矛盾.

情况 4：φ 形如 $\varphi'\rightarrow(x)\varphi'$,其中 x 在 φ' 中不出现, 假设存在 $\psi\in\mathrm{ST}(\mathscr{L})$ 使得 $P(\varphi,\ \neg\varphi\wedge\psi)=0$. 据定理 8.66,(1) $P(\varphi',\neg\varphi\wedge\psi)=1$. (2)$P((x)\varphi',\ \neg\varphi\wedge\psi)=0$. 由(2)和条件 8.18,得 $\lim\limits_{j\to\infty}P(\underbrace{((\cdots(\varphi'\wedge\varphi')\wedge\cdots)\wedge}_{j\text{ 项}}\varphi',\ \neg\varphi\wedge\psi)=0$,从而由定理 8.67 和极限定义得出,$P(\varphi',\ \neg\varphi\wedge\psi)=0$,这与(1)矛盾.

情况 5：φ 形如 $(x)\varphi'\rightarrow\varphi'(t/x)$. 假设存在 $\psi\in\mathrm{ST}(\mathscr{L})$ 使得 $P(\varphi,\ \neg\varphi\wedge\psi)=0$, 则据定理 8.66, 我们有 (1)$P((x)\varphi',\neg\varphi\wedge\psi)=1$. (2) 对$\mathscr{L}$的某个常元 t,$P(\varphi'$

(t/x), $\neg\varphi\wedge\psi) = 0$. 但据(1)和条件 8.18 得出,(3)

$$\lim_{j\to\infty} P\left(\bigwedge_{i=1}^{j} \varphi'(t_i/x),\ \neg\varphi\wedge\psi\right) = 1.$$

由定理 8.64 可知对从 1 开始的每一 j,

$$P\left(\bigwedge_{i=1}^{j} \varphi'(t_i/x),\ \neg\varphi\wedge\psi\right) = 1.$$

因此据定理 8.65，对 $\mathscr{L}$ 的每一常元 t，$P(\varphi'(t/x),\ \neg\varphi\wedge\psi) = 1$,这与(2)矛盾.

情况 6：φ 形如 $(x)(\varphi'\to\psi')\to((x)\varphi'\to(x)\psi')$.假设存在 $\psi\in \mathrm{ST}(\mathscr{L})$ 使得 $P(\varphi,\ \neg\varphi\wedge\psi) = 0$，则据定理 8.66 我们有，(1) $P((x)(\varphi'\to\psi'),\ \neg\varphi\wedge\psi) = 1$. (2) $P((x)\varphi'\to(x)\psi',\ \neg\varphi\wedge\psi) = 0$. 由(2)和定理 8.66 我们有,(3) $P((x)\varphi',\ \neg\varphi\wedge\psi) = 1$. (4) $P((x)\psi',\ \neg\varphi\wedge\psi) = 0$. 因此据(1),(3)和条件 8.18,有(5)

$$\lim_{j\to\infty} P\left(\bigwedge_{i=1}^{j} (\varphi'\to\psi')(t_i/x),\ \neg\varphi\wedge\psi\right) = 1,$$

和 (6)

$$\lim_{j\to\infty} P\left(\bigwedge_{i=1}^{j} \varphi'(t_i/x),\ \neg\varphi\wedge\psi\right) = 1.$$

由定理 8.64 我们有(7),对 $\mathscr{L}$ 的每一常元 t,

$$P((\varphi'\to\psi')(t/x),\ \neg\varphi\wedge\psi) = 1.$$

(8)对 $\mathscr{L}$ 的每一常元 t，$P(\varphi'(t/x), \neg\varphi\wedge\psi) = 1$. 由定理 8.69,(9)对 $\mathscr{L}$ 的每一常元 t，$P(\psi'(t/x), \neg\varphi\wedge\psi) = 1$. 因此据定理 8.70 和极限定义我们可得,(10)

$$\lim_{j\to\infty} P\left(\bigwedge_{i=1}^{j} \psi'(t_i/x),\ \neg\varphi\wedge\psi\right) = 1,$$

据条件 8.18，$P((x)\psi', \neg\varphi\wedge\psi) = 1$,这与(4)矛盾.

定理 8.72 令 φ 是 $\mathscr{L}$ 的公理，则对每一 $\psi\in \mathrm{ST}(\mathscr{L})$,

$P(\varphi,\psi)=1$.

据定理 8.71,8.69 以及极限定义和条件 8.18,可证定理成立。

定理 8.73 若 $S\vdash\varphi$，则对任意 $\psi\in \mathrm{ST}(\mathscr{L})$，若每一 $\theta\in S$，$P(\theta,\psi)=1$，则 $P(\varphi,\psi)=1$。

证明类似定理 8.42。

因为在定理 8.73 中用“P^+”置换“P”仍成立，所以我们有下列对应定理 8.43 的重要定理：

定理 8.74 若 $S\vdash\varphi$,则在二元概率意义上 S 逻辑推出 φ.

为了证明定理 8.74 的逆定理，还需以下定义和定理。

定义 8.7 令 $S\subseteq \mathrm{ST}(\mathscr{L})$，称函数 P_S 是 $\mathscr{L}$ 中 S 的二元相伴概率,若 P_S 满足下列条件：对任意 $\varphi,\psi\in \mathrm{ST}(\mathscr{L})$，若 $S\vdash\psi\to\varphi$，则 $P_S(\varphi,\psi)=1$；否则，$P_S(\varphi,\psi)=0$。

定理 8.75 令 $S\subseteq \mathrm{ST}(\mathscr{L})$，$P_S$ 是 $\mathscr{L}$ 中的二元相伴概率，则有

(1) P_S 满足条件 8.11—8.13,8.15—8.17。

(2) 若 S 在 $\mathscr{L}$ 中极大一致,则 P_S 满足条件 8.11—8.17。

(3) 若 S 在 $\mathscr{L}$ 中 ω完全，则 P_S 满足条件 8.11—8.13，8.15—8.18。

(4) 若 S 在 $\mathscr{L}$ 中极大一致且 ω完全，则 P_S 构成 $\mathscr{L}$ 的二元概率函数(即 P_S 满足条件 8.11—8.18)。

证明：(1)除了条件 8.15，其余的都可以从 P_S 的定义和 $\mathscr{L}$ 中关于证明论的基本事实得到。因此在此只须证 P_S 满足条件 8.15。

不失一般性假设 θ 是 P 正规的，且 $P_S(\varphi\wedge\psi,,\theta)=1$. 据定义 8.7，$S\vdash\theta\to\varphi\wedge\psi$,因此 $S\vdash\theta\to\psi$，$S\vdash\psi\wedge\theta\to\varphi$. 由定义 8.7，$P_S(\psi,\theta)=P_S(\varphi,\psi\wedge\theta)=1$,因此

$$P_S(\varphi\wedge\psi,\theta)=P_S(\psi,\theta)\cdot P_S(\varphi,\psi\wedge\theta),$$

再假设 $P_S(\varphi\wedge\psi,\theta)=0$，据定义 8.7，$S\nvdash\theta\to\varphi\wedge\psi$，因此有 $S\nvdash\theta\to\psi$ 或 $S\nvdash\psi\wedge\theta\to\varphi$。据定义 8.7，$P_S(\psi,\theta)=0$ 或 $P_S(\varphi,\psi\wedge\theta)=0$，所以仍有

$$P_S(\varphi\wedge\psi,\theta)=P_S(\psi,\theta)\cdot P_S(\varphi,\psi\wedge\theta).$$

(2) 假设 S 在 $\mathscr{L}$ 中极大一致且存在 $\theta\in ST(\mathscr{L})$ 使得 $P_S(\theta,\psi)\neq 1$。据定义 8.7，$S\nvdash\psi\to\theta$，因为 S 是极大一致的，所以 $\psi\to\theta\notin S$，$S\vdash\neg(\psi\to\theta)$。即(i) $S\vdash\psi\wedge\neg\theta$，我们先假设 $P_S(\neg\varphi,\psi)=1$，据定义 8.7，$S\vdash\psi\to\neg\varphi$。若 $S\vdash\psi\to\varphi$，则 $S\vdash\psi\to\neg\varphi\wedge\varphi$，$S\vdash\neg\psi$。再据(i)，就有 $S\vdash\psi\wedge\neg\psi\wedge\neg\theta$，这引起矛盾，所以 $S\nvdash\psi\to\varphi$。据定义 8.7，$P_S(\varphi,\psi)=0$。再假设 $P_S(\neg\varphi,\psi)=0$，则 $S\nvdash\psi\to\neg\varphi$。据 S 的极大一致性，$S\vdash\neg(\psi\to\neg\varphi)$。若 $S\nvdash\psi\to\varphi$，也有类似上面的矛盾，因此 $S\vdash\psi\to\varphi$，$P_S(\varphi,\psi)=1$。

(3) 证明类似定理 8.44(3)的证明。

定理 8.76 令 $S\subseteq ST(\mathscr{L})$ 且在 $\mathscr{L}$ 中极大一致，P_S 是 $\mathscr{L}$ 中 S 的二元相伴概率，$\varphi,\psi\in ST(\mathscr{L})$，则

(1) $\qquad\psi\to\varphi\in S\Longleftrightarrow P_S(\varphi,\psi)=1$.

(2)若 $\varphi\in S$，则 $P_S(\varphi,\neg(\psi\wedge\neg\psi))=1$。若 $\neg\varphi\in S$，则 $P_S(\varphi,\neg(\psi\wedge\neg\psi))=0$。

证明：(1) 若 $\psi\to\varphi\in S$，则 $S\vdash\psi\to\varphi$，据定义 8.7，$P_S(\varphi,\psi)=1$。若 $\psi\to\varphi\notin S$，则据 S 的极大一致性，$S\nvdash\psi\to\varphi$，据定义 8.7，$P_S(\varphi,\psi)=0$，因此 $P_S(\varphi,\psi)\neq 1$。
(2)若$\varphi\in S$，则 $\neg(\psi\wedge\neg\psi)\to\varphi\in S$，$S\vdash\neg(\psi\wedge\neg\psi)\to\varphi$，所以$P_S(\varphi,\neg(\psi\wedge\neg\psi))=1$。若$\neg\varphi\in S$，则 $\varphi\notin S$，$\neg(\psi\wedge\neg\psi)\to\varphi\notin S$，因此 $S\nvdash\neg(\psi\wedge\neg\psi)\to\varphi$，$P_S(\varphi,\neg(\psi\wedge\neg\psi))=0$。

定理 8.77 令 $S\subseteq ST(\mathscr{L})$ 且在 $\mathscr{L}$ 中无穷可扩张，$\varphi\in ST(\mathscr{L})$，若 $S\nvdash\varphi$，则存在 $\mathscr{L}$ 的二元概率函数 P 和 $\psi\in$

$\mathrm{ST}(\mathscr{L})$ 使得对每一 $\theta\in S$，$P(\theta,\psi)=1$，但 $P(\varphi,\psi)\neq 1$.

证明：假设 $S\nvdash\varphi$，其中 S 在 $\mathscr{L}$ 中无穷可扩张的，$H(S\cup\{\neg\varphi\})$ 是 $\mathscr{L}$ 中 $S\cup\{\neg\varphi\}$ 的 Henkin 扩张，P 是 $\mathscr{L}$ 中 $H(S\cup\{\neg\varphi\})$ 的二元相伴概率. 对任意 $\psi'\in\mathrm{ST}(\mathscr{L})$，令 $\psi=\neg(\psi'\wedge\neg\psi')$，因为 $H(S\cup\{\neg\varphi\})$ 在 $\mathscr{L}$ 中极大一致，所以据定理 8.76(2)，对 ψ 和每一 $\theta\in S$，$P(\theta,\psi)=1$ 且 $P(\varphi,\psi)=0$. 又因为 $H(S\cup\{\neg\varphi\})$ 在 $\mathscr{L}$ 中 ω 完全且极大一致，所以据定理 8.75(2)，P 构成 $\mathscr{L}$ 的二元概率函数.

因为在定理 8.75，8.76 中用“P^{+}”置换“P”，这些定理仍成立，所以据定理 8.77，我们有

定理 8.78 令 S 和 φ 定义如定理 8.77，如果对 $\mathscr{L}$ 的任意二元概率函数 P 和 $\mathscr{L}$ 的任意陈述句 ψ，若对每一 $\theta\in S$，$P(\theta,\psi)=1$，则 $P(\varphi,\psi)=1$，那么 $S\vdash\varphi$.

若 S 在 $\mathscr{L}$ 中不是无穷可扩张的，我们用类似定理 8.47 后面那样的处理. 据定理 8.78，我们有

定理 8.79 令任意 $S\subseteq\mathrm{ST}(\mathscr{L})$，任意 $\varphi\in\mathrm{ST}(\mathscr{L})$，若在二元概率的意义上 S 逻辑推出 φ，则 $S\vdash\varphi$.

因此，再据定理 8.74 且令 $S=\varnothing$，我们有

定理 8.80 令任意 $\varphi\in\mathrm{ST}(\mathscr{L})$，$\vdash\varphi\Longleftrightarrow\varphi$ 在二元概率意义上逻辑真.

下面我们讨论一元概率函数和二元概率函数的关系. 令 $\top=\neg(Q(t_1)\wedge\neg Q(t_1))$，其中 Q 是 $\mathscr{L}$ 的一元谓词，t_1 是 $\mathscr{L}$ 的字母顺序排列的第一个常元.

定理 8.81 令 P 是 $\mathscr{L}$ 的任意二元概率函数，则有

(1) $P(\varphi,\psi\wedge\top)=P(\varphi,\psi)$.

(2) 若存在 $\varphi\in\mathrm{ST}(\mathscr{L})$ 是 P 正规的，则 $\top$ 是 P 正规的.

(3) $\top$ 是 P 正规的.

证明：(1)据条件 8.15，$P(\top\wedge\varphi,\psi)=P(\top,\varphi\wedge\psi)\cdot$

$P(\varphi,\psi)$，由定理8.80，$P(\top,\varphi\wedge\psi)=1$，因此$P(\top\wedge\varphi,\psi)=P(\varphi,\psi)$，同理 $P(\varphi\wedge\top,\psi)=P(\varphi,\top\wedge\psi)\cdot P(\top,\psi)=P(\varphi,\top\wedge\psi)$. 因此由条件 8.16，$P(\varphi,\top\wedge\psi)=P(\varphi,\psi)$，由条件 8.17，$P(\varphi,\psi\wedge\top)=P(\varphi,\psi)$.

(2) 假设对任意 $\varphi\in ST(\mathscr{L})$，$P(\varphi,\top)=1$. 且令$\psi$是$\mathscr{L}$ 的任意陈述句，则⊤是 P 反常的，所以 $P(\varphi\wedge\psi,\top)=1$. 由条件 8.15，$P(\varphi,\psi\wedge\top)\cdot P(\psi,\top)=1$；由条件 8.12 和定理 8.58，$P(\varphi,\psi\wedge\top)=1$，由(1)，$P(\varphi,\psi)=1$. 因此，对任意 $\varphi,\psi\in ST(\mathscr{L})$，$P(\varphi,\psi)=1$. 据逆否命题，若存在 $\varphi,\psi\in ST(\mathscr{L})$，$P(\varphi,\psi)\neq 1$，则存在 $\varphi\in ST(\mathscr{L})$，

$$P(\varphi,\top)\neq 1.$$

(3) 据条件 8.11 和(2)，可证(3)成立.

值得注意，若我们令上述⊤是 $\mathscr{L}$ 的任意公理，仍有定理 8.81 成立，因此 $\mathscr{L}$ 的所有公理是 P 正规的.

定义 8.8 称二元概率函数 P 是对⊤的限制，若 P 的后一变目只取⊤为值. 称二元概率函数 P 有一个一元概率函数 $P_{\top}$ 作为 P 对⊤的限制，若 $P_{\top}(\varphi)=P(\varphi,\top)$ 且 $P_{\top}$ 满足条件 8.1—8.7. 称一元概率函数 $P_{\top}$ 是二元概率函数 P 对⊤的限制，若对每一 $\varphi\in ST(\mathscr{L})$，$P_{\top}(\varphi)=P(\varphi,\top)$.

定理 8.82 (1) $\mathscr{L}$ 的每一二元概率函数都有一个 $\mathscr{L}$ 的一元概率函数作为它的限制.

(2) $\mathscr{L}$ 的每一一元概率函数都是 $\mathscr{L}$ 的一个二元概率函数对⊤的限制.

证明：(1)令 $P_{\top}(\varphi)=P(\varphi,\top)$，下面证 $P_{\top}$ 满足条件 8.1—8.7. 据定理 8.81(3)和 8.68，

$$P(\varphi,\top)=P(\varphi\wedge\psi,\top)+P(\varphi\wedge\neg\psi,\top),$$

因此，$P_{\top}(\varphi)=P_{\top}(\varphi\wedge\psi)+P_{\top}(\varphi\wedge\neg\psi)$，而这就是条件 8.3. 利用条件 8.12，8.16 和 8.18，和上述方法直接证明 $P_{\top}$ 满

足条件 8.1，8.5 和 8.7。利用定理 8.73 和 8.67 和上述方法可证明 $P_{\top}$ 满足条件 8.2 和 8.4。通过重复使用条件 8.15 和定理 8.81(1)，可证明 $P_{\top}$ 满足条件 8.6。

(2) 令 P 是 $\mathscr{L}$ 的任意一元概率函数，定义 P' 如下：若 $P(\psi) \neq 0$，则 $P'(\varphi,\psi) = P(\varphi \wedge \psi)/P(\psi)$，否则

$$P'(\varphi,\psi) = 1.$$

令 $P'_{\top}$ 是 P' 对⊤的限制，则易证下列成立：(i) $P'_{\top}$ 构成 $\mathscr{L}$ 的二元概率函数。(ii)对任意 $\varphi \in \mathrm{ST}(\mathscr{L})$，

$$P'_{\top}(\varphi) = P(\varphi).$$

因为据定义，$P'_{\top}(\varphi) = P'(\varphi, \top)$，而据条件 8.2，$P(\top) \neq 0$。因此据上述定义，$P'_{\top}(\varphi) = P(\varphi \wedge \top)/P(\top)$，而由定理 8.10 和条件 8.2，$P(\varphi \wedge \top)/P(\top) = P(\varphi)$。因此存在 $\mathscr{L}$ 的二元概率函数 $P'_{\top}$ 使得 P 是 $P'_{\top}$ 对⊤的限制。

Leblanc 把概率作为认知主体的合理相信度赋于语言 $\mathscr{L}$ 的陈述句集，这是在条件 8.1—8.7，8.11—8.18(它们只刻划了一元和二元概率函数的纯形式特征）之外增加了非形式的语义信息，如果我们摒弃了这一部分非形式语义信息，Leblanc 在本章所得的结果仍然成立。不仅如此，上述结果还可以加强。例如对公理和推理规则还可以放宽，使它们也可以作用在公式上，而不仅仅象现在那样作用在句子上。这就说明，建立一个纯形式概率语义系统不仅是可能的，而且还是有用的。

第九章　无穷概率逻辑

在本世纪 60 年代 D. Scott 和 P. Krauss 在 [18] 一文在无穷概率逻辑方面做了许多开创性的工作。他们运用了布尔代数，测度论和模型论等工具，研究如何把概率赋于允许无穷长的公式，并初步建立了无穷概率逻辑的模型论，但当时他们还没有真正建立形式化的无穷概率逻辑的语法系统. 70 年代，D. N. Hoover 和 H. J. Keisler 建立了较完备的无穷概率逻辑的语法系统和模型论[19][20]。而后，Hoover 和 H. J. Keisler 等人逐步丰富了无穷概率逻辑，建立了带概率量词，随机变项算子，积分算子和条件期望算子的不同的无穷概率逻辑和内容丰富的模型论。无穷概率逻辑成为经典一阶逻辑的无穷扩张中最富成果的应用逻辑。

本章我们主要介绍 Model-Theoretic Logics 一书中 Keisler 对无穷概率逻辑的研究[21]，由于篇幅限制，我们在此只讨论带概率量词的无穷概率逻辑。

§1　具有概率量词的逻辑 $\mathscr{L}_{AP}$

本节我们引入逻辑 $\mathscr{L}_{AP}$，它非常类似于一阶无穷逻辑 $\mathscr{L}_A = A \cap \mathscr{L}_{\omega_1\omega}$，其中 A 是可容集，$\mathscr{L}_{\omega_1\omega}$ 是在普通带等词的一阶逻辑基础上添加可数合取号. 公理增加：$\wedge \Phi \rightarrow \varphi$，若 $\varphi \in \Phi$ 且 Φ 可数且 Φ 中只有有穷多个自由变元. 规则增加：

$$\{\psi \rightarrow \varphi : \varphi \in \Phi\} \vdash \psi \rightarrow \wedge \Phi^{[22]}.$$

逻辑 $\mathscr{L}_{AP}$ 没有普通的全称量词(x)和存在量词 $(\exists x)$，代之

以概率量词 $(px \geqslant r)$. 适合于这种逻辑的模型是一阶模型，并且它带有论域上的(可数可加)概率测度使每一关系可测. 公式

$$(px \geqslant r)\varphi(x)$$

表示集合 $\{x: \varphi(x)\}$ 至少有概率 r. 这一节我们提出逻辑 $\mathscr{L}_{AP}$ 的公理、推理规则和关于概率模型的定义以及基本概念,以便后几节更详细地研究和讨论.

1. 语法

本章一直假设 A 是可容集[22],并且对于 $\omega \in$ A,每一 $a \in$ A 是可数的,即 A$\subseteq$HC,其中 HC 是遗传可数集的集合.

今后一直假设语言 $\mathscr{L}$ 是由有穷元的关系符号和个体常元符号(除非另加说明)组成的 A 递归集，注意 $\mathscr{L}$ 中没有函数符号. 在[19]中，Hoover 建立的无穷概率逻辑的语言有函数符号.

定义 9.1 逻辑 $\mathscr{L}_{AP}$ 有下列逻辑符号:

(1) 可数个个体变元 x_n, 对 $n \in$ N (自然数集).

(2) 联结词$\neg$和$\wedge$.

(3) 量词$(P\bar{x} \geqslant r)$, 其中 $\bar{x} = \langle x_1, \cdots, x_n\rangle$ 是不同的个体变元的一个 n 元组, $r \in$ A$\cap[0,1]$.

(4) 等词符号=.

定义 9.2 $\mathscr{L}_{AP}$ 的公式集是使得下列条件成立的最小的集合

(1) 一阶逻辑的每一原子公式是 $\mathscr{L}_{AP}$ 的公式.

(2) 若φ是 $\mathscr{L}_{AP}$ 的公式,则 $\neg\varphi$ 也是 $\mathscr{L}_{AP}$ 的公式.

(3) 若 $\Phi \in$ A 是 $\mathscr{L}_{AP}$ 的公式集且只有有穷多个自由变元,则$\wedge\Phi$ 是 $\mathscr{L}_{AP}$ 的公式.

(4) 若φ是 $\mathscr{L}_{AP}$ 的公式且 $(P\bar{x} \geqslant r)$ 是 $\mathscr{L}_{AP}$ 的概率量词,则 $(P\bar{x} \geqslant r)\varphi$ 是 $\mathscr{L}_{AP}$ 的公式.

注意：上述公式能集合论地构造起来使得 $\mathscr{L}_{AP}$ 的公式集是 A 的子集．若我们用 $\mathscr{L}_{\omega_1 P}$ 指称当 A = HC 时的 $\mathscr{L}_{AP}$，则

$$\mathscr{L}_{AP} = A \cap \mathscr{L}_{\omega_1 P}.$$

自由变元和约束变元的概念定义如通常，量词 $(P\bar{x} \geqslant r)$ 约束 n 元组 $\bar{x}$ 中所有的个体变元、等词关系在逻辑 $\mathscr{L}_{AP}$ 中只起较小的作用，这是因为缺少全称量词和存在量词的缘故．

定义 9.3 在 $\mathscr{L}_{AP}$ 中引入下列约定．

(1) $(P\bar{x} < r)\varphi \triangleq \neg(P\bar{x} \geqslant r)\varphi$.

(2) $(P\bar{x} \leqslant r)\varphi \triangleq (P\bar{x} \geqslant 1 - r)\neg\varphi$.

(3) $(P\bar{x} > r)\varphi \triangleq \neg(P\bar{x} \geqslant 1 - r)\neg\varphi$.

(4) $\bigvee_{\varphi \in \Phi} \varphi \triangleq \neg \bigwedge_{\varphi \in \Phi} \neg\varphi$.

(5) $\rightarrow, \leftrightarrow$ 定义如通常．

从直观上说，量词 $(P\bar{x} \geqslant 1)$ 类似一阶逻辑中的全称量词 $(\bar{x})$，但比 $(\bar{x})$ 要弱；而量词 $(P\bar{x} > 0)$ 类似一阶逻辑中的存在量词 $(\exists\bar{x})$，但比 $(\exists\bar{x})$ 要强． 原则上可以只利用约束一个变元的概率量词 $(Px \geqslant r)$，而把 $(P\bar{x} \geqslant r)$ 作为缩写引入，但这样的缩写相当复杂，所以较简单的做法是在语言中直接引入 $(P\bar{x} \geqslant r)$．

2. 概率模型

在定义概率模型之前，我们先来介绍概率论和测度论的某些非常基本的概念．首先，一个有穷可加概率空间是 $\langle A, S, \mu\rangle$，其中 S 是 A 的子集的域，μ 是 S 上的实值函数 $\mu: S \to [0,1]$，并且使得 $\mu(A) = 1$，且对 $X, Y \in S$，有

$$\mu(X \cup Y) = \mu(X - Y) + \mu(Y - X) + \mu(X \cap Y).$$

我们称集合 $X \in S$ 为 μ 可测的，μ 称为 A 上的有穷可加概率测度，$\langle A, S, \mu\rangle$ 称为概率空间．若 S 是 σ 域且 μ 是可数可加

的，即在 S 中若 $X_0 \subseteq X_1 \subseteq \cdots$，则

$$\mu\left(\bigcup_n^{\infty} X_n\right) = \lim_{n \to \infty} \mu(X_n).$$

在这种情况下，μ 称为 A 上的可数可加概率测度。为了简单起见，我们以后就用"概率测度"来表示"（可数可加）概率测度"。

集合 X 称为 μ 的零测集，若存在 $Y \supseteq X$ 使得 $\mu(Y) = 0$。两个概率空间 $\langle A, S, \mu\rangle$ 和 $\langle B, T, \nu\rangle$ 的积是下列概率空间：

$$\langle A \times B, S \otimes T, \mu \otimes \nu\rangle,$$

其中 $A \times B$ 是笛卡尔积，$S \otimes T$ 是由可测矩形集 $X \times Y$（其中 $X \in S, Y \in T$）生成的 σ 代数，且

$$(\mu \otimes \nu)(X \times Y) = \mu(X) \cdot \nu(Y).$$

上述乘积空间是二重乘积空间，我们也可简单地用 $\langle A^n, S^n, \mu^n\rangle$ 来指称 n 重乘积空间。

一般说，对角集 $\{\langle x, x\rangle : x \in A\}$ 不是 μ^2 可测的。但是若每一单元集是可测的，则存在一种典范方式把积测度扩张到这样的对角集上。在每一单元集有零测度的情况下，这样的对角集也有零测度。但一般说来，只有可数多个单元集有正测度，并且这样的对角集的测度是这些单元集测度的平方和。

定义 9.4 令 $\langle A, S, \mu\rangle$ 是概率空间，并且使得每一单元集可测，则对每一 $n \in \mathrm{N}$，$\langle A^n, S^{(n)}, \mu^{(n)}\rangle$ 是概率空间，并且使得 $S^{(n)}$ 是 S^n 的可测矩形和对角集

$$D_{ij} = \{\bar{x} \in A^n : x_i = x_j\}$$

生成的 σ 代数，且 $\mu^{(n)}$ 是 μ^n 对 $S^{(n)}$ 的唯一扩张

$$\mu^{(n)}(D_{ij}) = \sum_{x \in A} (\mu(\{x\}))^2.$$

定理 9.1 若 $\langle A, S, \mu\rangle$ 是使得每一单元集可测的概率空间，则由定义 9.4 给出的 $S^{(n)}$ 的测度 $\mu^{(n)}$ 存在且唯一。因此

对每一 $x \in S^{(n)}$，存在一 μ^n 可测集 U 使得 $\mu^{(n)}(X \triangle U) = 0$，其中 $X \triangle U$ 表示 X 和 U 的对称差：

$$X \triangle U = (X - U) \cup (U - X).$$

证明：我们在此给出 $n = 2$ 的证明．因为

$$S^{(2)} = \{X \subseteq A^2 : X = (Y \cap D_{12}) \cap (Z - D_{12}),$$

$$\text{其中 } Y, Z \in S^2\}.$$

定义 $\nu: S^{(2)} \to [0,1]$ 如下：

$$\nu(X) = \sum_{\langle x,x\rangle \in Y} (\mu(\{x\}))^2 + \mu^2(Z) - \sum_{\langle x,x\rangle \in Z} (\mu(\{x\}))^2,$$

显然 $\nu = \mu^{(2)}$ 是 $S^{(2)}$ 上唯一的可数可加概率测度，且扩张 μ^2 满足

$$\nu(D_{12}) = \sum_{x \in A} (\mu(\{x\}))^2.$$

令 $E = \{\langle x,x\rangle \in D_{12} : \mu(\{x\}) > 0\}$，$U = (Y \cap E) \cup (Z - E)$，因为 E 可数且每一单元集可测，所以 E，从而 U 是 μ^2 可测的．又因为 $\nu(D_{12} - E) = 0$ 且 $x \triangle U \subseteq D_{12} - E$，所以我们有 $\nu(X \triangle U) = 0$．证毕．

现在我们定义语言 $\mathscr{L}$ 的概率模型．

定义 9.5 $\mathscr{L}$ 的概率模型为

$$\mathfrak{U} = \langle A, R_i^{\mathfrak{U}}, c_j^{\mathfrak{U}}, \mu\rangle_{i \in I, j \in J},$$

其中 $\langle A, R_i^{\mathfrak{U}}, c_j^{\mathfrak{U}}\rangle_{i \in I, j \in J}$ 是经典一阶模型，μ 是 A 上的（可数可加）概率测度，并且使得每一单元集可测，任一 $R_i^{\mathfrak{U}}$ 是 $\mu^{(n)}$ 可测的．

因此我们对 $\mathfrak{U} \models \varphi[\bar{a}]$（即在 $\mathfrak{U}$ 中 $\bar{a} \in A^n$ 满足 $\varphi(\bar{x})$，其中 $\bar{x} = \langle x_1, \cdots, x_n\rangle$）可以给出下列递归定义：

(1) $\mathfrak{U} \models Q[\bar{a}] \Longleftrightarrow Q(\bar{x})$ 是 $\mathscr{L}_{AP}$ 的原子公式，且 $\bar{a} \in Q^{\mathfrak{U}}$．

(2) $\mathfrak{U} \models \neg\varphi[\bar{a}] \Longleftrightarrow \mathfrak{U} \not\models \varphi[\bar{a}]$，即 $\bar{a}$ 在 $\mathfrak{U}$ 中不满足

$\varphi(\bar{x})$。

（3） $\mathfrak{U}\models\bigwedge_{\varphi\in\Phi}\varphi[\bar{a}]\Longleftrightarrow$ 对每一 $\varphi\in\Phi,\mathfrak{U}\models\varphi[\bar{a}]$，其中 $\Phi\in\mathrm{A}$ 是 $\mathscr{L}_{\mathrm{A}P}$ 的可数公式集。

（4） $\mathfrak{U}\models(P\bar{y}\geqslant r)\varphi(\bar{x},\bar{y})[\bar{a}]\Longleftrightarrow\{\bar{b}\in A^m:\mathfrak{U}\models\varphi[\bar{a},\bar{b}]\}$ 是 $\mu^{(m)}$ 可测的且至少有测度 r。

定理 9.2 对每一概率模型 $\mathfrak{U}$，以及每一公式 $\varphi(\bar{x},\bar{y})\in\mathscr{L}_{\mathrm{A}P}$ 和 A 中每一序列 $\bar{a}$，集合 $\{\bar{b}\in A^n:\mathfrak{U}\models\varphi[\bar{a},\bar{b}]\}$ 是 $\mu^{(n)}$ 可测的。

我们利用定理 9.2 证明 $\mathscr{L}_{\mathrm{A}P}$ 的满足关系中有我们想要的意义。该定理的证明可以利用下面的 Fubini 定理的"对角"形式和数学归纳法得到。

定理 9.3 (Fubini 定理) 令 μ 是概率测度使得每一单元集可测，$B\subseteq A^{m+n}$ 是 $\mu^{(m+n)}$ 可测的，则

（1） 每一节 $B_{\bar{x}}=\{\bar{y}\in A^n:\bar{x}\bar{y}\in B\}$ 是 $\mu^{(n)}$ 可测的。

（2） 函数 $f(\bar{x})=\mu^{(n)}(B_{\bar{x}})$ 是 $\mu^{(m)}$ 可测的。

（3） $\mu^{(m+n)}(B)=\int f(\bar{x})d\mu^{(m)}$。

证明：该定理的证明与通常的积测度的 Fubini 定理的证明相同。若我们想在语言中同时引入全称量词和概率量词，则定理 9.3 不成立，这是因为可测集的投影自身无需是可测的。

模型论的概念：同构，$\mathscr{L}_{\mathrm{A}P}$ 初等等价，$\mathscr{L}_{\mathrm{A}P}$ 初等子模型定义如通常，且分别记为 $\mathfrak{U}\cong\mathfrak{B},\mathfrak{U}\equiv_{\mathrm{A}P}\mathfrak{B},\mathfrak{U}\prec_{\mathrm{A}P}\mathfrak{B}$。

3. 证明论

Keisler 的带概率量词的无穷概率逻辑有两个公理化语法系统，一个称为弱 $\mathscr{L}_{\mathrm{A}P}$，一个称为完备 $\mathscr{L}_{\mathrm{A}P}$。

定义 9.6 弱 $\mathscr{L}_{\mathrm{A}P}$ 的公理(模式)和推理规则如下：其中的公式和公式集都是 $\mathscr{L}_{\mathrm{A}P}$ 的公式和公式集，实数 r，$s\in$

$A\cap[0,1]$.

公理 9.1 $\mathscr{L}_{A}$ 的所有无量词公理:

(1) 古典命题逻辑的重言式及其代入特例。

(2) $\bigwedge\Phi\rightarrow\varphi$, 其中 $\varphi\in\Phi$.

(3) $x=x$.

(4) $x=y\rightarrow y=x$.

(5) $\varphi(x)\wedge t=x\rightarrow\varphi(t)$, 其中 $\varphi(x)$ 是公式, t 是项, $\varphi(t)$ 是用 t 置换 $\varphi(x)$ 每一自由出现的 x 得到的公式。

公理 9.2 (单调性)

$$(P\bar{x}\geqslant r)\varphi\rightarrow(P\bar{x}\geqslant s)\varphi,\ \text{其中}\ r\geqslant s.$$

公理 9.3 $(P\bar{x}\geqslant r)\varphi(\bar{x})\rightarrow(P\bar{y}\geqslant r)\varphi(\bar{y})$.

公理 9.4 $(P\bar{x}\geqslant 0)\varphi$.

公理 9.5 (有穷可加性)

(1) $(P\bar{x}\leqslant r)\varphi\wedge(P\bar{x}\leqslant s)\psi\rightarrow((P\bar{x}\leqslant r+s)(\varphi\vee\psi))$.

(2) $(P\bar{x}\geqslant r)\varphi\wedge(P\bar{x}\geqslant s)\psi\wedge(P\bar{x}\leqslant 0)(\varphi\wedge\psi)\rightarrow$
$(P\bar{x}\geqslant r+s)(\varphi\vee\psi)$.

公理 9.6 (Archimedean 性质)

$$(P\bar{x}>r)\varphi\leftrightarrow\bigvee_{n\in\mathbb{N}}(P\bar{x}\geqslant r+1/n)\varphi.$$

推理规则:

规则 9.1 $\varphi,\varphi\rightarrow\psi\vdash\psi$.

规则 9.2 (合取规则)

$$\{\psi\rightarrow\varphi:\varphi\in\Phi\}\vdash\psi\rightarrow\bigwedge\Phi.$$

规则 9.3 (概括规则)

$\varphi\rightarrow\psi(\bar{x})\vdash\varphi\rightarrow(P\bar{x}\geqslant 1)\psi(\bar{x})$, 其中 $\bar{x}$ 在 φ 中不自由.

定义 9.7 完备 $\mathscr{L}_{AP}$ 是在弱 $\mathscr{L}_{AP}$ 系统上再加下列公理.

公理 9.7 (可数可加性)

$$\bigwedge_{\substack{\Psi\subseteq\Phi\\ f}} (P\bar{x} \geqslant r)\wedge\Psi \to (P\bar{x} \geqslant r)\wedge\Phi,$$

其中Ψ跑遍Φ的有穷子集.

公理 9.8（对称性）

$$(Px_1\cdots x_n \geqslant r)\varphi \leftrightarrow (Px_{\pi 1}\cdots x_{\pi n} \geqslant r)\varphi,$$

其中π是$\{1,\cdots,n\}$的一置换（permutation）.

此公理说明个体变元的位置不影响φ的成立与否.

公理 9.9（乘积独立性） 若$\bar{x},\bar{y}$中的变元都不同，则

$$(P\bar{x} \geqslant r)(P\bar{y} \geqslant s)\varphi \to (P\bar{x}\bar{y} \geqslant r\cdot s)\varphi.$$

公理 9.10（乘积可测性） 对每一$r<1$，若$\bar{x},\bar{y}$中的变元都不同，则

$$(P\bar{x} \geqslant 1)(P\bar{y} > 0)(P\bar{z} \geqslant r)(\varphi(\bar{x},\bar{z}) \leftrightarrow \varphi(\bar{y},\bar{z})).$$

公理 9.10 的中心目的是要确保可以用可测矩形的有穷并来近似$\varphi(\bar{x},\bar{y})$. 若$\varphi(\bar{x},\bar{z})$是“矩形”$\psi(\bar{x})\wedge\theta(\bar{z})$，公理 9.10 显然是有效的. 我们在后面能看到在一般情况下这一公理也是有效的(可靠性定理).

定理 9.4（演绎定理） 在完备$\mathscr{L}_{AP}$或弱$\mathscr{L}_{AP}$中，若ψ是一句子且$\Phi\cup\{\psi\}\vdash\theta$，则$\Phi\vdash\psi\to\theta$.

证明：首先证明

$$\vdash \bigwedge_{\varphi\in\Phi}(\psi\to\varphi)\to(\varphi\to\wedge\Phi). \tag{9.1}$$

据公理 9.1(2)，有$\bigwedge_{\varphi\in\Phi}(\psi\to\varphi)\to(\psi\to\varphi)$.

从而$\bigwedge_{\varphi\in\Phi}(\psi\to\varphi)\wedge\psi\to\varphi$.

因此由规则 9.2 得$\bigwedge_{\varphi\in\Phi}(\psi\to\varphi)\wedge\psi\to\bigwedge_{\varphi\in\Phi}\varphi$,

$$\bigwedge_{\varphi\in\Phi}(\psi\to\varphi)\to(\psi\to\wedge\Phi).$$

现在我们对复杂度用归纳法来证明演绎定理。令 θ_1，θ_2，$\cdots$是从 $\Phi\cup\{\phi\}$ 得到的一个证明使得该证明序列的最后一个公式是 θ（注意该证明序列可以是无数无穷的，这一点和有穷逻辑不一样）。对任意 $i=1,2,\cdots$，当 θ_i 是公理，或 $\theta_i\in\Phi$，或 θ_i 是 ϕ，或 θ_i 是通过规则 9.1 从该证明序列更前面的公式得到时，证明如通常。非平凡的两步是考虑规则 9.2 和 9.3：

情况 1．假设 θ_i 是从该证明序列更前面的若干公式 $\theta_{i_k}=\varphi\rightarrow\gamma_{i_k}$（其中 $j_k<i$）根据规则 9.2 得到的，即 $\theta_i=\varphi\rightarrow\bigwedge\Gamma$（其中 $\Gamma=\{\gamma_{i_k}:\theta_i$ 从 $\varphi\rightarrow\gamma_{i_k}$ 通过规则 9.2 得到$\}$)，则据归纳假设，对每一 $\gamma_{i_k}\in\Gamma$，$\Phi\vdash\phi\rightarrow(\varphi\rightarrow\gamma_{i_k})$。由规则 9.2，$\Phi\vdash\phi\rightarrow\bigwedge_{\gamma_{i_k}\in\Gamma}(\varphi\rightarrow\gamma_{i_k})$，再由(9.1)，$\Phi\vdash\phi\rightarrow(\varphi\rightarrow\bigwedge\Gamma)$，即 $\Phi\vdash\phi\rightarrow\theta_i$。

情况 2．假设 θ_i 从该证明序列的更前的某个公式 $\theta_j=\varphi\rightarrow\gamma(\bar{x})$（其中 $j<i$）根据规则 9.3 得到的，即 $\theta_i=\varphi\rightarrow(P\bar{x}\geqslant 1)\gamma(\bar{x})$，且 $\bar{x}$ 在 φ 中不自由出现。据归纳假设，$\Phi\vdash\phi\rightarrow(\varphi\rightarrow\gamma(\bar{x}))$，其中 $\bar{x}$ 在 φ 中不自由出现，则 $\Phi\vdash\phi\wedge\varphi\rightarrow\gamma(\bar{x})$。因为 $\bar{x}$ 在 $\phi\wedge\varphi$ 中不自由出现（因为 ϕ 是句子），所以据规则 9.3，$\Phi\vdash\phi\wedge\varphi\rightarrow(P\bar{x}\geqslant 1)\gamma(\bar{x})$，即 $\Phi\vdash\phi\rightarrow(\varphi\rightarrow(P\bar{x}\geqslant 1)\gamma(\bar{x}))$，因此 $\Phi\vdash\phi\rightarrow\theta_i$。证毕。

定理 9.5 下列定理是完备 $\mathscr{L}_{AP}$ 的定理且它们的证明无需用到公理 9 10。

(1) $(P\bar{x}\leqslant 1)\varphi$.

(2) $(P\bar{x}>r)\bigvee\Phi\leftrightarrow\bigvee_{\psi\in\Phi}(P\bar{x}>r)\bigvee\Psi$.

(3) $(P\bar{x}\geqslant r)\varphi\leftrightarrow\bigwedge_{n\in\mathbb{N}}(P\bar{x}\geqslant r-1/n)\varphi$.

(4) $(P\bar{x}\leqslant r)\varphi(\bar{x})\wedge(P\bar{y}\leqslant s)\phi(\bar{y})\rightarrow(P\bar{x}\bar{y}\leqslant r\cdot s)(\varphi$

$(\bar{x}) \wedge \psi(\bar{y}))$.

(5) $(P\bar{x} \geqslant r)\varphi(\bar{x}) \to (P\bar{x}\bar{y} \geqslant r)\varphi(\bar{x})$.

(6) $(P\bar{x}\bar{y} \geqslant r + s - r \cdot s)\varphi \to (P\bar{x} \geqslant r)(P\bar{y} \geqslant s)\varphi$.

(7) $(P\bar{x} \geqslant 1)(P\bar{y} \geqslant 1)\varphi \leftrightarrow (P\bar{x}\bar{y} \geqslant 1)\varphi$.

定理 9.6 完备 $\mathscr{L}_{AP}$ 的任意有概率模型的句子集是一致的。

象通常一样只须证明每一公理是有效的，且推理规则保持有效性。唯一的困难是检验公理 9.10 的有效性。根据定理 9.5(3)，只须证明对每一$q,r < 1$，

$$(P\bar{x} \geqslant q)(P\bar{y} > 0)(P\bar{z} \geqslant r)(\varphi(\bar{x}\bar{z}) \leftrightarrow \varphi(\bar{y}\bar{z}))$$

是有效的。而这可用 Fubini 定理，定理 9.1 和下列定理(2)⇒(1)的方向来证明。

定理 9.7 令 μ, ν 和 λ 分别是 M, N 和 $M \times N$ 上的概率测度，且 $\mu \otimes \nu \subseteq \lambda$，$U$ 是 λ 可测的，则下列(1)和(2)等价。

(1) 对每一$\varepsilon > 0$，存在 $\mu \otimes \nu$ 可测矩形的有穷并B使得 $\lambda(U \triangle B) < \varepsilon$。

(2) 存在一 $\mu \otimes \nu$ 可测集C使得 $\lambda(U \triangle C) = 0$。

证明：(1)⇒(2)，我们可以取C作为这样的B的极限。(2)⇒(1)，可以用单调类定理证明，对每一 $\mu \otimes \nu$ 可测的 U，(1)成立。这立即得到(2)⇒(1)。

4. 弱概率模型

定义 9.8 完备 $\mathscr{L}_{AP}$ 的弱概率模型是下列结构

$$\mathfrak{U} = \langle A, R_i^{\mathfrak{U}}, c_j^{\mathfrak{U}}, \mu_n \rangle_{i \in I, j \in J, n \in N},$$

其中每一 μ_n 是 A^n 上的有穷可加概率，使得每一单元集可测，且对每一 $\varphi(\bar{x}, \bar{y}) \in \mathscr{L}_{AP}$ 和A中的每一序列 $\bar{a}$，集合$\{\bar{b} \in A^n: \mathfrak{U} \models \varphi[\bar{a}, \bar{b}]\}$ 是 μ_n 可测的。

根据定理 9.2，每一概率模型 $\mathfrak{U}$ 导出 $\mathscr{L}_{AP}$ 的一个弱概率模型使得 $\mu_n = \mu^{(n)}$。

定理 9.8（弱可靠性定理） 令Φ是语言$\mathscr{L}_{AP}$的句子集，若Φ有一弱概率模型，则Φ在弱$\mathscr{L}_{AP}$中一致。

定理 9.9（弱完全性定理） 令 A 可数，若Φ在弱$\mathscr{L}_{AP}$中一致，则Φ有一弱概率模型。

定理可以证明如下：令C是新个体常元集，语言$K=\mathscr{L}\cup C$。据 Henkin 构造，Φ能扩张到具有下列性质的极大弱K_{AP}一致句子集Γ：

(1) 若 $\Phi\subseteq\Gamma$ 且 $\wedge\Phi\in K_{AP}$，则 $\wedge\Phi\in\Gamma$。

(2) 若对所有C中序列 $\bar{c}$，$\varphi(\bar{c})\in\Gamma$，则 $(P\bar{x}\geqslant r)\varphi(\bar{x})\in\Gamma$。

C'是语言K的常元集。Γ用通常的构造等价类的方法（即定义等价关系$\approx$：$c\approx c'\Longleftrightarrow c=c\in\Gamma$）导出$K$的一个一阶模型

$$\mathfrak{U}_0=\langle A,R_i^*,c^{\mathfrak{U}}\rangle_{i\in I,c\in C'},$$

其中 $A=\{c^{\mathfrak{U}}:c\in C'\}$，$c^{\mathfrak{U}}$实际上是$c$的等价类 $c/\approx$。对每一$\varphi(\bar{x},\bar{y})$ 和$\bar{d}$，定义

$$\mu_n\{\bar{c}^{\mathfrak{U}}:\varphi(\bar{c},\bar{d})\in\Gamma\}=\sup\{r:(P\bar{x}\geqslant r)\varphi(\bar{x},\bar{d})\in\Gamma\}.$$

公理 9.1—9.5 确保 μ_n 是良定义的，且是有穷可加概率测度。这就给出一个弱概率模型。公理 9.6（用定理 9.5(3) 的对偶形式）确保上述上确界总能得到。根据数学归纳法，$\mathfrak{U}\models\Gamma$，因此 $\mathfrak{U}\models\Phi$。这就证明了定理。

5. 原子概率模型和可数概率模型

假设$\mathscr{L}$没有个体常元符号。

定义 9.9 令$\mathfrak{U}$是概率模型。称一元素$a\in A$是原子，若$\{a\}$有正测度。称$\mathfrak{U}$是原子模型，若A中每一元素是原子。

定理 9.10 (1) 若$\mathfrak{U}$是原子概率模型，则$\mathfrak{U}$是可数概率模型。

(2) 若$\mathfrak{U}$是可数概率模型，则$\mathfrak{U}$满足

$$(Px \geqslant 1)(Py > 0)x = y. \tag{9.2}$$

(3) 若 $\mathfrak{U}$ 满足(9.2)，则存在唯一的原子概率模型使得 $\mathfrak{B} \prec_{AP} \mathfrak{U}$.

(4) 若 $\mathfrak{U}$ 和 $\mathfrak{B}$ 是原子概率模型且 $\mathscr{L}_{AP}$ 等价，则它们同构。

(5) 若 $\mathfrak{U}$ 是原子概率模型，则对每一公式 $\varphi(x, \bar{y}) \in \mathscr{L}_{AP}$ 和 A 中的每一序列 $\bar{b}$，我们有

$$\mathfrak{U} \models (x)\varphi(x,\bar{b}) \longleftrightarrow (Px \geqslant 1)\varphi(x,\bar{b}),$$

$$\mathfrak{U} \models (\exists x)\varphi(x,\bar{b}) \longleftrightarrow (Px > 0)\varphi(x,\bar{b}),$$

其中，对于全称量词和存在量词，$\mathfrak{U}$ 看作标准一阶模型 $\mathfrak{U}_0$，(5)说明了，在原子概率模型中，可以用概率量词来定义普通量词。

定理 9.11（原子模型的完全性定理） $\mathscr{L}_{AP}$ 的可数句子集 Φ 有一原子模型 $\Leftrightarrow \Phi \cup \{(9.2)\}$ 在 $\mathscr{L}_{AP}$ 中一致。

证明： 令 A 可数。假设 $\Phi \cup \{(9.2)\}$ 在 $\mathscr{L}_{AP}$ 中一致，则它有一弱概率模型 $\mathfrak{U}_0$ 使得 $\mathscr{L}_{AP}$ 的所有定理在 $\mathfrak{U}_0$ 中成立。由完备 $\mathscr{L}_{AP}$，对每一m，可以从(9.2)中导出

$$(Px \geqslant 1)\bigvee_n (Py \geqslant 1/n)x = y.$$

$$(Px > 1 + 1/m)\bigvee_n (Py \geqslant 1/n)x = y,$$

$$\bigvee_n (Px > 1 - 1/m)(Py \geqslant 1/n)x = y.$$

因此 $\mathfrak{U}_0$ 有 μ_1 测度小于 $1-1/m$ 的有穷子集，用下列等式把 μ_1 扩张到定义在 A_0 的所有子集上的概率测度 μ

$$\mu(X) = \sup\{\mu_1(Y): Y \subseteq_f X\},$$

从而形成一个概率模型 $\mathfrak{U}(\equiv_{AP}\mathfrak{U}_0)$. 据定理 9.10(3)，原子模型 $\mathfrak{B}(\prec_{AP}\mathfrak{U})$ 就是Φ的原子模型。 另一方向的证明据定理

9.6. 证毕。

§2 完全性定理

这一节将证明§1给出的公理集是完全的。作为初步的结果，我们在本节2证明完备 $\mathscr{L}_{AP}$ —{公理 9.10} 的完全性定理。在完全性证明中的主要困难是从有穷可加概率模型构造可数可加概率模型。克服这个困难的关键是用非标准分析构造 Loeb 测度。

1. Loeb 测度

以下我们一直假设有一个 ω_1 饱和的非标准论域

$$*:\langle V_\omega(U), \in\rangle \to \langle V_\omega(^*U), \in\rangle,$$

其中 U 是足够大的始元 (urelement) 集。对 $r\in {}^*\mathrm{R}$, $^\circ r$ 表示 r 的标准部分。

定义9.10 令 B 是 $V_\omega(^*U)$ 中一个内集，$\langle B,S,\mu\rangle$ 是一个 $*$ 有穷的可加概率空间(因此 μ 和 S 都是内的且 $\mu: S\to {}^*[0,1]$)。μ 的 Loeb 测度是使得下列条件成立的唯一的(可数可加)概率空间 $\langle B, \sigma(S), \hat{\mu}\rangle$:

(1) $\sigma(S)$ 是由 S 生成的 σ 代数。

(2) 对所有 $X\in S$, $\hat{\mu}(X)={}^\circ\mu(X)$.

定理 9.12 上述 Loeb 测度存在且唯一。

可以利用 ω_1 饱和定理和 Caratheodory 扩张定理来证明。

定理 9.13 令 $X\in\sigma(S)$，则

(1) 对每一 $n\in\mathrm{N}$, 存在 Y, $Z\in S$ 使得 $Y\subseteq X\subseteq Z$, 且 $\mu(Z-Y)<1/n$.

(2) 存在 $Y\in S$ 使得 $\hat{\mu}(X\triangle Y)=0$.

利用单调类定理和 ω_1 饱和定理来证明。

2. 分级概率模型

定义 9.11 $\mathscr{L}$ 的分级概率模型为

$$\mathfrak{U}=\langle A,R_i^{\mathfrak{U}},c_j^{\mathfrak{U}},\mu_n\rangle_{i\in I,j\in J,n\in\mathbb{N}},$$

并且使得

(1) 每一 μ_n 是 A^n 上的(可数可加)概率测度.

(2) 每一 n 元关系 $R_i^{\mathfrak{U}}$ 是 μ_n 可测的,等词关系是 μ_2 可测的.

(3) 若 A' 是 μ_m 可测的,则 $A'\times A^n$ 是 μ_{m+n} 可测的.

(4) 对称性成立,即每一 μ_n 在 $\{1,\cdots,n\}$ 的置换下保持不变.

(5) $\langle\mu_n:n\in\mathrm{N}\rangle$ 具有 Fubini 性质,即若 A' 是 μ_{m+n} 可测的,则 (i) 对每一 $\bar{x}\in A^m$, 节 $A'_{\bar{x}}=\{\bar{y}:A'(\bar{x},\bar{y})\}$ 是 μ_n 可测的. (ii) 函数 $f(\bar{x})=\mu_n(A'_{\bar{x}})$ 是 μ_m 可测的. (iii) $\int f(\bar{x})d_{\mu_m}=\mu_{m+n}(A')$.

定理 9.14 (1) 若 $\mathfrak{U}$ 是概率模型,则

$$\langle A,R_i^{\mathfrak{U}},c_j^{\mathfrak{U}},\mu^{(n)}\rangle_{i\in I,j\in J,n\in\mathrm{N}}$$

是一分级概率模型.

(2) 每一分级概率模型是 $\mathscr{L}_{\omega_1 P}$ 的弱模型.

定理 9.15 在分级概率模型 $\mathfrak{U}$ 中, μ_n 是 $\mu_1^{(n)}$ 的扩张.

定理 9.16 令 A 是*有穷集, 对每一 n, 令 ν_n 是 A^n 上的内概率测度,并且使得给每一元素以相同的权(计数测度), $\mu_n=\hat{\nu}_n$ 是 ν_n 的 Loeb 测度, 则 $\langle\mu_n:n\in\mathrm{N}\rangle$ 有 Fubini 性质. 因此,若 A 的每一 n 元关系是 μ_n 可测的,则 $\mathfrak{U}$ 是分级概率模型.

定义 9.12 所谓分级(逻辑) $\mathscr{L}_{AP}$ 是指完备逻辑 $\mathscr{L}_{AP}$-{公理 9.10}

定理 9.17 (分级可靠性定理) 有分级模型的 $\mathscr{L}_{AP}$ 的句

子集在分级 $\mathscr{L}_{AP}$ 中是一致的.

定理 9.18（分级完全性定理） 每一在分级 $\mathscr{L}_{AP}$ 中一致的可数句子集 Φ 有分级模型.

证明：令 A 可数，且假设 $\mathscr{L}$ 有可数个不在 Φ 中出现的常元. 根据弱完全性定理的证明，Φ 有弱概率模型：

$$\mathfrak{U}=\langle A, R_i^{\mathfrak{U}}, c_j^{\mathfrak{U}}, \mu_n\rangle_{i\in I, j\in J, n\in\mathbf{N}},$$

并且 $\mathfrak{U}$ 满足分级 $\mathscr{L}_{AP}$ 的每一定理，$A=\{c_j: j\in J\}$. 每一 μ_n 的定义域是 A^n 的 $\mathscr{L}_{AP}$ 可定义的子集的集合，因此形成内概率模型

$$^*\mathfrak{U}=\langle {}^*A, {}^*R_i^{\mathfrak{U}}, {}^*c_j^{\mathfrak{U}}, {}^*\mu_n\rangle_{i\in {}^*I, j\in {}^*J, n\in {}^*\mathbf{N}}.$$

令

$$\hat{\mathfrak{U}}=\langle {}^*A, {}^*R_i^{\mathfrak{U}}, {}^*c_j^{\mathfrak{U}}, \hat{\mu}_n\rangle_{i\in I, j\in J, n\in\mathbf{N}},$$

其中 $\hat{\mu}_n$ 是 μ_n 的 Loeb 测度. 据定理 9.13，每一 $\hat{\mu}_n$ 可测集用 n 个变元的 $*$ 可定义集上下近似. 从而利用公理9.8—9.9，可以证明 $\hat{\mathfrak{U}}$ 是分级概率模型. 运用归纳法于公式的复杂度能证明 $\hat{\mathfrak{U}}\equiv \mathbf{A}_P\mathfrak{U}$. 其中，对 $\wedge$ 步骤用公理 9.7，对量词步骤用公理 9.8—9.9，因此 $\hat{\mathfrak{U}}\models\Phi$.

3. 主要完全性结果

定理 9.19 令 $\mathfrak{A}$ 是满足 $\mathscr{L}_{AP}$ 的每一定理的分级概率模型，则对每一 $\varepsilon>0$ 和每一公式 $\varphi(\bar{y})\in\mathscr{L}_{AP}$，存在有穷多个公式 $\psi_{ij}(\bar{x}, y)(i=1,\cdots,m, j=1,\cdots,n)$ 使得

$$\mathfrak{U}\models(P\bar{x}>0)(P\bar{y}>1-\varepsilon)\Big(\varphi(\bar{y})\longleftrightarrow\bigvee_{i=1}^{m}\bigwedge_{j=1}^{n}\psi_{ij}(\bar{x}, y_j)\Big).$$

这个定理断定 $\mathfrak{U}$ 中任一可定义集 $\varphi(\bar{y})$ 可以用可定义矩形的有穷并在 ε 范围内近似.

要反复利用公理 9.10 来证明这定理.

定义 9.13 令 $\mathfrak{U}$ 和 $\mathfrak{B}$ 是分级概率模型，若 $\mathfrak{U}$ 和 $\mathfrak{B}$ 有相同的论域，常元和测度，且对每一公式 $\varphi(\bar{x})\in\mathscr{L}_{AP}$ 和 μ_n 的几

乎所有 $\bar{a}$，有

$$\mathfrak{A}\models\varphi[\bar{a}]\Longleftrightarrow\mathfrak{B}\models\varphi[\bar{a}],$$

则称 $\mathfrak{A}=\mathfrak{B}$ 几乎确定 (almost surely)，记作 $\mathfrak{A}=\mathfrak{B}$ a. s.

定理 9.20 令 $\mathfrak{A}=\mathfrak{B}$ a. s.，则 $\mathfrak{A}\equiv_{\mathrm{AP}}\mathfrak{B}$，而且对每一 $\varphi(\bar{x})\in\mathscr{L}_{\mathrm{AP}}$ 和 μ_n 的几乎所有 $\bar{a}$

$$\mathfrak{A}\models\varphi[\bar{a}]\Longleftrightarrow\mathfrak{B}\models\varphi[\bar{a}].$$

利用归纳法可以证明上述定理成立.

定理 9.21 令 $\mathfrak{A}$ 是满足 $\mathscr{L}_{\mathrm{AP}}$ 的每一定理的分级概率模型,则存在一分级概率模型 $\mathfrak{B}$ 使得 $\mathfrak{A}=\mathfrak{B}$ a. s.，且每一关系 $R_i^{\mathfrak{B}}$ 是 $\mu^{(n)}$ 可测的,因此 $\mathfrak{B}$ 导出一个普通概率模型.

证明：据矩形近似定理，对每一 $\varepsilon>0$，和 $\mathfrak{A}$ 中每一 $\mathscr{L}_{\mathrm{AP}}$ 可定义集 $U\subseteq A^n$，存在 μ^n 可测矩形的有穷并使得 $\mu_n(B\Delta U)<\varepsilon$. 因此据定理 9.7，存在 μ^n 可测集 C 使得 $\mu_n(C\Delta U)=0$. 通过补缀上对角集,对每一 $i\in I$，存在 $\mu^{(n)}$ 可测关系 $R_i^{\mathfrak{B}}$ 使得 $\mathfrak{A}=\mathfrak{B}$ a. s.，定理得证.

定理 9.22 ($\mathscr{L}_{\mathrm{AP}}$ 的完全性定理) $\mathscr{L}_{\mathrm{AP}}$ 的每一可数一致句子集 Φ 都有一个概率模型

利用定理 9.18，9.21 和 9.20 来证明.

4. 有穷概率逻辑

我们现在简单讨论 $\omega\notin\mathbf{A}$ 使得 $\mathscr{L}_{\mathrm{AP}}$ 的每一公式有穷的情况. 我们假设有理数以 $\mathbf{Q}\subseteq\mathbf{A}$ 的方式得到定义，因此 $\mathscr{L}_{\mathrm{AP}}$ 至少有量词 $(P\bar{x}\geqslant r)$，$r\in\mathbf{Q}\cap[0,1]$. 当 $\omega\notin\mathbf{A}$ 时，无穷公理 9.7 和无穷合取规则 9.2 退化为有穷情况,而另一个无穷公理 9.6 代之以新的无穷推理规则:

$$\{\psi\to(P\bar{x}\geqslant r)(P\bar{y}\geqslant s-1/n)\varphi:n\in\mathbb{N}\}$$
$$\vdash\psi\to(P\bar{x}\geqslant r)(P\bar{y}\geqslant s)\varphi.$$

据上述新增规则,弱概率模型,分级概率模型和完全性定理对

有穷情况 $\omega \notin A$ 都成立。

例如，令Φ是包含 $(Px>0)R(x)$ 和 $(Px\leqslant 1/n)R(x)$（其中 $n=1,2,\cdots$，R 是一元关系）的句子集，则Φ的每一有穷子集有模型，但Φ自身没有模型。

该例说明紧致性定理对 $\mathscr{L}_{AP}$ 一般不成立。

定义 9.14 $\mathscr{L}_{\omega_1 P}$的全称合取公式集是一最小公式集，它包含所有无量词公式，且在任意∧，有穷∨和量词 $(P\bar{x}\geqslant r)$ 下封闭

定理 9.23（有穷紧致性定理） 令Φ是 $\mathscr{L}_{AP}$ 的全称合取句子集．若Φ的每一有穷子集有分级概率模型，则Φ有分级概率模型。

证明：假设每一 $\Psi\subseteq_f\Phi$ 有一模型 $\mathfrak{U}_\Psi$，取所有 $\mathfrak{U}_\Psi$ 的超积模型 $^*\mathfrak{U}$ 使得对每一 $\varphi\in\Phi$，几乎每一 $\mathfrak{U}_\Psi$ 满足 φ。据 Loeb 的构造，从 $^*\mathfrak{U}$ 可以构造一分级概率模型 $\hat{\mathfrak{U}}$，于是据数学归纳法证明，在几乎所有 $\mathfrak{U}_\Psi$ 中成立的每一全称合取公式也在 $\hat{\mathfrak{U}}$ 中成立，证毕.

5. $\mathscr{L}_{\omega_1\omega}$ 的句子上的概率

$\mathscr{L}_{\omega_1\omega}$ 的句子是 $\mathscr{L}_{\omega\omega}$（普通一阶逻辑）的句子在可数合取号上的扩张．在此，我们讨论两种量词都出现的逻辑。

定义 9.15 $\mathscr{L}_A(P,\forall)$ 是两种量词逻辑：一种是概率量词 $(P\bar{x}\geqslant r)$，用 $x,y,z,\cdots$ 表示这种量词中的个体变元，另一种是全称量词(t)，用 $s,t,\cdots$表示这种量词中的个体变元。$\mathscr{L}_A(P,\forall)$ 的概率模型有下列形式．

$$\mathfrak{U}=\langle A,T,R_i,c_j,\mu\rangle_{i\in I,j\in J},$$

其中μ是A上的概率测度，且对每一 $\bar{t}\in T$，$R_i(\bar{x};\bar{t})$ 是 $\mu^{(n)}$ 可测的。

若T可数，则在$\mathfrak{U}$中定义包括通常的全称量词(t)的满足

关系是不困难的。还有一种满足关系的定义能运用于任意的 $\mathscr{L}_{\mathrm{A}}(P,\forall)$ 的概率模型。D. Scott 和 P. Krauss 在[18]中作了类似的叙述：对每一 $\varphi(\bar{x};\bar{t})$ 和每一 $\bar{b}\in T$，指派模为零测集的 $\mu^{(n)}$ 的测度代数的一个元素 $\varphi(\bar{x};\bar{b})^{\mathfrak{u}}$。因此全称量词的满足关系为

$$(t)\varphi(\bar{x};\bar{b})^{\mathfrak{u}}=\inf\{\varphi(\bar{x};\bar{b},c)^{\mathfrak{u}}:c\in T\},$$

其中 inf 是取该测度代数的下确界。这是和当 T 可数时满足关系的自然定义相符的，但和 T 不可数时不相符。

定义 9.16 $\mathscr{L}_{\mathrm{A}}(P,\forall)$ 公理(模式)由 $\mathscr{L}_{\mathrm{AP}}$ 的公理和下列量词公理组成。

公理 9.11 $(t)\varphi(t)\to\varphi(t')$，其中 t' 是项，$\varphi(t')$ 是用 t' 置换 $\varphi(t)$ 中所有自由出现的 t 得到的。

公理 9.12

$$(P\bar{x}\geqslant r)(t)\varphi(\bar{x};t)\longleftrightarrow$$

$$\bigwedge_{n}(t_1)\cdots(t_n)(P\bar{x}\geqslant r)\bigwedge_{k=1}^{n}\varphi(\bar{x};t_k),$$

其中公理 9.11 是经典一阶逻辑的量词公理。公理 9.12 是新增公理。

$\mathscr{L}_{\mathrm{A}}(P,\forall)$ 的推理规则是 $\mathscr{L}_{\mathrm{AP}}$ 的推理规则，再加上下列经典一阶逻辑的全称概括规则。

规则 9.4 $\psi\to\varphi(t)\vdash\psi\to(t)\varphi(t)$，其中 t 在 ψ 中不自由出现。

定理 9.24（可靠性定理） $\mathscr{L}_{\mathrm{AP}}(P,\forall)$ 的每一有概率模型的句子集 Φ 都是一致的。

定理 9.25（完全性定理） $\mathscr{L}_{\mathrm{AP}}(P,\forall)$ 的每一可数一致句子集都有一 T 可数的概率模型 $\mathfrak{u}$。

我们可以构造一个可数弱概率模型，然后固定第二种量

词，并且运用本节2和3的方法把第一种量词扩张到一个概率模型来证明本定理。

定义 9.17 $\mathscr{L}$ 逻辑是把概率赋于 $\mathscr{L}_{\omega_1\omega}=\mathscr{L}_{HC}$ 的句子的逻辑(它的模型论已在 [18] 中讨论过了)。

$\mathscr{L}$ 概率逻辑的概率模型为

$$\mathfrak{A}=\langle A, T, R_i, \mu\rangle.$$

$\mathfrak{A}$ 可以看作是语言 $\mathscr{L}$ 的一阶模型 $\mathfrak{A}_x$ 的指标族 $\{\mathfrak{A}_x: x\in A\}$，即每一 $\mathfrak{A}_x$ 的论域都是 T，且具有 A 上的概率测度 μ 使得每一 $\{x: R_i(x;\bar{t})\}$ 可测。

定义 9.18 $\mathscr{L}_{\omega_1\omega}$ 上的概率是从 $\mathscr{L}_{\omega_1\omega}$ 的句子集到[0,1]的函数，对于 $\vee$，$\neg$，它是可数可加的，且每一有效句子的测度为1。

从 $\mathscr{L}$ 逻辑的每一个概率模型 $\mathfrak{A}$ 导出在 $\mathscr{L}_{\omega_1\omega}$ 上的概率 $\mu^{\mathfrak{A}}$

$$\mu^{\mathfrak{A}}(\varphi)\geqslant r \Longleftrightarrow \mathfrak{A}\models(Px\geqslant r)\varphi.$$

$\mathscr{L}$ 概率逻辑的公理和推理规则是 $\mathscr{L}_{HC}(P,\forall)$ 的公理和推理规则，但除去公理 9.3，9.8—9.10。显然可靠性和完全性定理仍成立。下列的完全性定理在 [18] 中证明了。

定理 9.26 令 μ 是 $\mathscr{L}_{\omega_1\omega}$ 上对纯等词句赋于1或0的概率。对每一可数集 $\Phi\subseteq\mathscr{L}_{\omega_1\omega}$，存在 $\mathscr{L}$ 概率逻辑的概率模型 $\mathfrak{A}$ 使得 $\mu^{\mathfrak{A}}$ 和 Φ 上的 μ 相同。

§3 模 型 论

在本节我们将建立逻辑 $\mathscr{L}_{AP}$ 的模型论。其中叙述了大数定律的模型论形式，它说明每一概率模型都可以用几乎所有的有穷子概率模型来"近似"。此外还用这些结果来证明超

穷模型的存在性和唯一性，$\mathscr{L}_{AP}$ 的 Robinson 一致性定理和 Craig 内插定理。

1. 大数定律

本小节的结果对所有的分级概率模型都成立。

定义 9.19 $\mathscr{L}_{AP}$ 的有穷全称公式为

$$(P\bar{x}_1 \geqslant r_1)\cdots(P\bar{x}_n \geqslant r_n)\phi,$$

其中 ϕ 是 $\mathscr{L}$ 的有穷无量词公式。

$\mathscr{L}_{AP}$ 的有穷存在公式为

$$(P\bar{x}_1 > r_1)\cdots(P\bar{x}_n > r_n)\phi,$$

其中 ϕ 是 $\mathscr{L}$ 的有穷无量词公式。

注意：因为 $\neg(P\bar{x} > r)\phi$ 等价 $(P\bar{x} \geqslant 1 - r)\neg\phi$，所以有穷存在(全称)公式的否定等价有穷全称(存在)公式。

定义 9.20 令 $\mathfrak{A}$ 是 $\mathscr{L}$ 的分级概率模型，$\bar{a}_K = \langle a_1, \cdots, a_K\rangle \in A^K$ 是 A 的 K 元组，有穷样本 $\mathfrak{A}(\bar{a}_K)$ 是一概率模型，它的论域是 $\{a_1, \cdots, a_K\}$ 和 $\mathfrak{A}$ 的常元(若有的话)的并，$\mathfrak{A}(\bar{a}_K)$ 的一阶部分是 $\mathfrak{A}$ 的子模型，且它的测度由下列等式给出

$$\nu(S) = |\{m \leqslant K : a_m \in S\}|/K,$$

因此在 $\mathfrak{A}(\bar{a}_K)$ 中有穷集 $\{a_1, \cdots, a_K\}$ 有测度 1，且单元集 $\{a\}$ 的测度是 $1/K$ 乘以序列 $\bar{a}_K$ 中 a 出现的次数。

定理 9.27 令 $\mathfrak{A}$ 是 $\mathscr{L}$ 的具有测度 μ_n 的分级概率模型，φ 是 $\mathscr{L}_{AP}$ 的有穷存在句，且 $\mathfrak{A} \models \varphi$，则有

(1) $\mathscr{L}_{AP}$ 的弱大数定律

$$\lim_{K\to\infty} \mu_K\{\bar{a}_K \in A^K : \mathfrak{A}(\bar{a}_K) \models \varphi\} = 1.$$

(2) $\mathscr{L}_{AP}$ 的强大数定律：令 $\mu_{\mathbf{N}}$ 是由 μ_n 确定的 $A^{\mathbf{N}}$ (其中 N 是自然数集)上的测度的完备化，则对 $\mu_{\mathbf{N}}$ 的几乎所有序列 $\bar{a} \in A^{\mathbf{N}}$，以及对所有有穷的 $K \in \mathbf{N}$，$\mathfrak{A}(\bar{a}_K) \models \varphi$.

定理 9.28 (范式定理) 分级 $\mathscr{L}_{\omega_1 P}$ 的每一公式 $\varphi(\bar{x})$ 等价于形如 $(P\bar{y} \geqslant r)\phi(\bar{x}, \bar{y})$ 的公式的可数布尔组合 (即只

用 $\neg$和$\wedge$的组合)，其中 $\phi(\bar{x},\bar{y})$ 是$\mathscr{L}$ 的原子公式的有穷合取。

证明：根据前束范式论证能说明分级 $\mathscr{L}_{\omega_1 P}$ 的每一公式等价有穷全称公式（有相同自由变项）的一个可数布尔组合。根据弱大数定律，下面每一陈述蕴涵下一个陈述：其中ϕ是一个有穷无量词公式。

(1) $\mathfrak{A}\models(P\bar{x}\geqslant r)(P\bar{y}\geqslant s)\phi$.

(2) $\bigwedge_n \mathfrak{A}\models(P\bar{x}>r-1/n)(P\bar{y}>s-1/n)\phi$.

(3) $\bigwedge_n \lim_{K\to\infty}\mu_K\{\bar{a}_K:\mathfrak{A}(\bar{a}_K)\models(P\bar{x}>r-1/n)(P\bar{y}>s-1/n)\phi\}=1$.

(4) $\bigwedge_n \lim_{K\to\infty}\mu_K\{\bar{a}_K:\mathfrak{A}(\bar{a}_K)\models(P\bar{x}\geqslant r-1/n)(P\bar{y}\geqslant s-1/n)\phi\}>0$.

(5) $\bigwedge_n \mathfrak{A}\models(P\bar{x}\geqslant r-1/n)(P\bar{y}\geqslant s-1/n)\phi$.

(6) $\mathfrak{A}\models(P\bar{x}\geqslant r)(P\bar{y}\geqslant s)\phi$.

因此

$$\mathfrak{A}(\bar{z}_K)\models(P\bar{x}\geqslant r-1/n)(P\bar{y}\geqslant s-1/n)\phi$$

在$\mathfrak{A}$中可以用 $\mathscr{L}$ 的一个有穷无量词公式 $\theta(\bar{z}_K)$ 来表达。分级 $\mathscr{L}_{\omega_1 P}$ 的每一公式等价于形如 $(P\bar{z}\geqslant t)\theta$ 的公式的可数布尔组合，其中θ是有穷无量词公式。最后我们用下列概率规则把θ归约为原子公式的合取：

$$P(\neg\varphi)=1-P(\varphi),$$

$$P(\varphi\vee\psi)=P(\varphi)+P(\psi)-P(\varphi\wedge\psi).$$

定理 9.29 令$\mathfrak{A}$和$\mathfrak{B}$是$\mathscr{L}$的分级概率模型，则下列陈述等价：

(1) $\mathfrak{A}\equiv_{\omega_1 P}\mathfrak{B}$.

(2) $\mathfrak{A}\equiv_{A P}\mathfrak{B}$.

(3) $\mathfrak{A}\models\varphi \Longleftrightarrow \mathfrak{B}\models\varphi$，其中$\varphi$是具有定理 9.28 范式形式的 $\mathscr{L}_{AP}$ 的任一句子。

下列结果是用有穷样本中的真值刻划$\equiv_{AP}$的性质。在普通的一阶逻辑模型论中没有对应的结果。

定理 9.30 令$\mathfrak{A}$和$\mathfrak{B}$是$\mathscr{L}$的分级概率模型，下列陈述等价：

(1) $\mathfrak{A}\equiv_{AP}\mathfrak{B}$.

(2) 对 $\mathscr{L}_{AP}$ 的每一句子φ和 $K\in N$,

$$\mu_K\{\bar{a}_K:\mathfrak{A}(\bar{a}_K)\models\varphi\}=\nu_K\{\bar{b}_K:\mathfrak{B}(\bar{b}_K)\models\varphi\},$$

即在$\mathfrak{A}$的K个元素的样本集中φ的概率和在$\mathfrak{B}$的K个元素的样本集中φ的概率相同。

(3) 对 $\mathscr{L}_{AP}$ 的每一句子φ,

$$\lim_{K\to\infty}\mu_K\{\bar{a}_K:\mathfrak{A}(\bar{a}_K)\models\varphi\}=1$$

$$\Longleftrightarrow \lim_{K\to\infty}\nu_K\{\bar{b}_K:\mathfrak{B}(\bar{b}_K)\models\varphi\}=1.$$

证明：利用范式定理，也可以利用 $\mathscr{L}_{AP}$ 的弱大数定律来直接证明。(1)⇒(2) 对每一K和ϕ，存在 $\mathscr{L}_{AP}$ 的一公式 $\phi(x_1,\cdots,x_K)$，它表示一个由K个元素组成的样本满足φ。(2)⇒(3) 显然。下证(3)⇒(1)。假设(3)成立，且令$\varphi(\bar{x})$ 是有穷无量词公式。据弱大数定律，我们有

$$\lim_{K\to\infty}\mu_K\{\bar{a}_K:\mathfrak{A}(\bar{a}_K)\models(P\bar{x}>s)\varphi\}=1.$$

据(3)，同样的结果也在$\mathfrak{B}$中成立。再用弱大数定律，有 $\mathfrak{B}\models(P\bar{x}\geqslant s)\varphi$。因为对所有$s<r$，上述成立，所以 $\mathfrak{B}\models(P\bar{x}\geqslant r)\varphi$，因此从范式定理得到 $\mathfrak{A}\equiv_{\omega_1P}\mathfrak{B}$.

2. 超穷模型

在本小节中我们假设$\mathscr{L}$只有有穷多个常元符号。

定义 9.21 (齐一)有穷概率模型是一概率模型$\mathfrak{A}$，其论域A有穷且测度是计数测度 $\mu(Y)=|Y|/|A|$. *有穷概率

模型是一在非标准论域上的有穷概率模型. 超穷概率模型是一概率模型 $\mathfrak{U}$, 其论域 A 是一无穷的 * 有穷集, 且 μ 是由 A 上的 * 计数测度确定的 Loeb 测度. 超穷分级概率模型是一分级概率模型, 其论域 A 是一无穷的 * 有穷集, 且每一 μ_n 是由 A^n 上的 * 计数测度确定的 Loeb 测度.

定理 9.31 每一超穷概率模型或分级模型是无原子的.

定理 9.31 是定理 9.16的重述.

定理 9.32 令 $\mathfrak{U}_0$ 是普通一阶模型, 其论域是一无穷的 * 有穷集, 且 $\mathfrak{U}_0$ 的每一关系是相对 A^n 上的计数测度的 Loeb 测度, 则存在唯一超穷分级概率模型以 $\mathfrak{U}_0$ 作为它的一阶部分.

现在我们引入超穷模型和 * 有穷模型之间的重要关系: 提升 (lifting) 关系.

定义 9.22 令 $\mathfrak{U}$ 是超穷分级概率模型, $\mathfrak{U}$ 的一个提升是一 * 有穷概率模型 $\mathfrak{B}$ 使得 $\mathfrak{B}$ 有和 $\mathfrak{U}$ 相同的论域和常元, 且对每一原子公式 $\varphi(\bar{x})$, 集合

$$\{\bar{a}:\mathfrak{U}\models\varphi[\bar{a}]\Longleftrightarrow\mathfrak{B}\models\varphi[\bar{a}]\}$$

有 μ_n 测度 1. 所谓超穷概率模型 $\mathfrak{U}$ 的提升是指与 $\mathfrak{U}$ 具有相同一阶部分的唯一的超穷分级概率模型 $\mathfrak{U}'$ 的提升.

定理 9.33 (1) 每一无穷的 * 有穷概率模型是某个超穷分级概率模型的提升.

(2) 每一超穷分级概率模型有一个提升.

(3) 若 $\mathfrak{U}$ 和 $\mathfrak{B}$ 是有一个共同提升的超穷分级概率模型, 则 $\mathfrak{U}=\mathfrak{B}$ a. s. 且 $\mathfrak{U}\equiv_{\omega_1 P}\mathfrak{B}$.

利用定理 9.32, 定理 9.13, 和定理 9.20, 可证明定理成立.

定理 9.34 (超穷概率模型存在定理) 令 $\mathfrak{B}$ 是 $\mathscr{L}$ 的无原子概率模型, A 是一无穷的 * 有穷集, 则存在超穷概率模型

$\mathfrak{A}$使得它的论域是 A，且 $\mathfrak{A}\equiv_{AP}\mathfrak{B}$.

证明：先假设 $\mathscr{L}$ 无常元符号，令 S 是 $\mathfrak{B}$ 的元素的所有无穷序列 $\bar{a}$ 的集合使得对 $\mathscr{L}_{AP}$ 的每一有穷存在公式 φ，所有有穷的 K，若 $\mathfrak{B}\models\varphi$，则 $\mathfrak{B}(\bar{a}_K)\models\varphi$，据弱大数定律，$\nu^{N}(S)=1$，因为 $\mathfrak{B}$ 是无原子的，所以 ν^{N} 的几乎每一序列 $\bar{a}$ 是一对一的，因此每一 $\mathfrak{B}(\bar{a}_K)$ 是齐一有穷概率模型. 由于存在 $\bar{a}\in S$ 使得 $\bar{a}$ 是一对一的，K 是一无穷超穷整数，则 $\mathfrak{B}(\bar{a}_K)$ 是 * 基数为 K 的 * 有穷概率模型且是超穷分级概率模型的提升. 因为对每一有穷无量词公式 $\phi(\bar{x})$ 和每一 r，$\bar{a}\in S$，所以 $\mathfrak{B}\models(P\bar{x}\geqslant r)\phi\Rightarrow\mathfrak{A}'\models(P\bar{x}\geqslant r)\phi$. 据范式定理，$\mathfrak{A}'\equiv_{AP}\mathfrak{B}$，由定理 9.21，存在一个超穷概率模型 $\mathfrak{A}$ 使得 $\mathfrak{A}=\mathfrak{A}$ a.s.，因此 $\mathfrak{A}\equiv_{AP}\mathfrak{B}$.

若 $\mathscr{L}$ 有有穷个常元，证明也相同，除了 $^{*}\mathfrak{B}(\bar{a}_K)$ 上的测度与计数测度有所不同，因为常元只有零测度. 证毕.

概率模型存在定理对分级模率模型也成立，证明也相同，对无等词的 $\mathscr{L}_{AP}$，定理无需假设$\mathfrak{B}$无原子也成立.

定义 9.23 令 $\mathfrak{A}$ 和 $\mathfrak{B}$ 是概率模型，从 $\mathfrak{A}$ 到 $\mathfrak{B}$ 的几乎确定的同构（用符号表示：$h:\mathfrak{A}\cong\mathfrak{B}$ a. s.）是双射 $h:\mathfrak{A}\to\mathfrak{B}$，且有

（1）h 是概率空间上的同构：$h:\langle A,\mu\rangle\cong\langle B,\nu\rangle$.

（2）对每一原子公式 $\varphi(\bar{x})$，在 $\mu^{(n)}$ 中下列等价关系几乎确定

$$\mathfrak{A}\models\varphi[\bar{a}]\Longleftrightarrow\mathfrak{B}\models\varphi[h\bar{a}].$$

定理 9.35 假设 $h:\mathfrak{A}\cong\mathfrak{B}$ a. s.，则

（1）对 $\mathscr{L}_{AP}$ 的每一公式 $\varphi(\bar{x})$，

$$\mathfrak{A}\models\varphi[\bar{a}]\Longleftrightarrow\mathfrak{B}\models\varphi[h\bar{a}]$$

在 $\mu^{(n)}$ 中几乎确定.

（2）$\mathfrak{A}\equiv_{AP}\mathfrak{B}$.

利用归纳法来证明.

定理 9.36（超穷概率模型的唯一性定理） 令 $\mathfrak{A}$ 和 $\mathfrak{B}$ 是有相同论域 A 的超穷概率模型，则下列条件是等价的.

(1) $\mathfrak{A}\equiv_{\mathrm{A}P}\mathfrak{B}$.

(2) 存在一 $h:\mathfrak{A}\cong\mathfrak{B}$ a. s..

(3) 存在一内 h 使得 $h:\mathfrak{A}\cong\mathfrak{B}$ a. s..

证明：这里我们只证明(1)⇒(3). 考虑任何内双射保持测度和 $\mathscr{L}$ 的原子公式的 n 元组： $\varphi_1(\bar{y}),\cdots,\varphi_n(\bar{y})$， 令 $\varepsilon>0$，据定理 9.19，可以找到一双射 $h_0:A\to A$ 使得

$$\mu^{(m)}\left(\bigcap_{K=1}^{n}\{\bar{a}:\mathfrak{A}\models\varphi_K[\bar{a}]\Longleftrightarrow\mathfrak{B}\models\varphi_K[h_0\bar{a}]\}\right)\geqslant1-\varepsilon.$$

这个思想就是用定理 9.13 来挑选一个 h_0 以近似保持这些矩形的每一坐标使得这些矩形近似 φ_K，现在令 $\hat{\mathfrak{A}}$，$\hat{\mathfrak{B}}$ 分别是 $\mathfrak{A}$，$\mathfrak{B}$ 的提升，据 ω_1 饱和性，我们能找到一内双射 h 使得对所有原子公式 $\varphi(\bar{y})$ 和所有实数 $\varepsilon>0$，

$$\mu^{(m)}(\{\bar{a}:\hat{\mathfrak{A}}\models\varphi[\bar{a}]\Longleftrightarrow\hat{\mathfrak{B}}\models\varphi[h\bar{a}]\})\geqslant1-\varepsilon,$$

由此可得 $h:\mathfrak{A}\cong\mathfrak{B}$ a. s.，因此(3)成立. 其余的方向易证.

3. Robinson 一致性定理与 Craig 内插定理

在证明这些定理时，超穷模型和饱和模型在普通一阶模型论中的作用相同.

定理 9.37（$\mathscr{L}_{\mathrm{A}P}$ 的 Robinson 一致性定理） 令 $\mathscr{L}^1$ 和 $\mathscr{L}^2$ 是两语言，$\mathscr{L}^0=\mathscr{L}^1\cap\mathscr{L}^2$，$\mathfrak{A}^1$ 和 $\mathfrak{A}^2$ 分别是 $\mathscr{L}^1$ 和 $\mathscr{L}^2$ 的概率模型，且它们的归约模型 $\mathfrak{A}^1\restriction\mathscr{L}^0\equiv_{\mathscr{L}^0_{\mathrm{A}P}}\mathfrak{A}^2\restriction\mathscr{L}^0$，则存在 $\mathscr{L}^1\cup\mathscr{L}^2$ 的概率模型 $\mathfrak{B}$ 使得

$$\mathfrak{B}\restriction\mathscr{L}^1\equiv_{\mathscr{L}^1_{\mathrm{A}P}}\mathfrak{A}^1,$$

$$\mathfrak{B}\restriction\mathscr{L}^2\equiv_{\mathscr{L}^2_{\mathrm{A}P}}\mathfrak{A}^2.$$

证明：我们给出 $\mathscr{L}^1$ 和 $\mathscr{L}^2$ 没有常元且 $\mathfrak{A}^1$ 和 $\mathfrak{A}^2$ 是无原

子概率模型时的证明（一般情况可以通过对原子公式增加新关系和扩张无原子部分得到）。据定理 9.34，我们可以把$\mathfrak{A}^1$和$\mathfrak{A}^2$看作是有相同论域A的超穷概率模型．据定理 9.36，存在一内双射 $h: \mathfrak{A}^1 \upharpoonright \mathscr{L}^0 \cong \mathfrak{A}^2 \upharpoonright \mathscr{L}^0$ a. s.．通过重新命名元素，我们可以把h当作恒等映射．通过换变$\mathscr{L}^0$的零测集上的$\mathfrak{A}^2$的关系，得到 $\mathfrak{A}^1 \upharpoonright \mathscr{L}^0 = \mathfrak{A}^2 \upharpoonright \mathscr{L}^0$ a. s.，令$\mathfrak{B}$是$\mathfrak{A}^1$和$\mathfrak{A}^2$的共同的膨胀概率模型，即$\mathfrak{B}$为所求的．证毕．

定理 9.38（$\mathscr{L}_{AP}$的 Craig 内插定理） 令 $\mathscr{L}^0 = \mathscr{L}^1 \cap \mathscr{L}^2$，$\varphi^1 \in \mathscr{L}^1_{AP}$，$\varphi^2 \in \mathscr{L}^2_{AP}$ 是两句子，使得$\models \varphi^1 \to \varphi^2$，则存在句子 $\varphi^0 \in \mathscr{L}^0_{AP}$ 使得 $\models \varphi^1 \to \varphi^0$ 且 $\models \varphi^0 \to \varphi^2$．

证明：假设没有这样的φ^0，据 Henkin 构造，存在$\mathscr{L}^1_{AP}$的φ^1的弱概率模型 $\mathfrak{A}^1$ 和$\mathscr{L}^2_{AP}$的φ^2的弱概率模型 $\mathfrak{A}^2$ 使得

$$\mathfrak{A}^1 \upharpoonright \mathscr{L}^0 \equiv_{\mathscr{L}^0_{AP}} \mathfrak{A}^2 \upharpoonright \mathscr{L}^0,$$

且 $\mathscr{L}^1_{AP}$ 和 $\mathscr{L}^2_{AP}$ 的公理皆有效．据完全性定理，我们得到φ^1的概率模型$\mathfrak{B}^1$和$\neg\varphi^2$的概率模型$\mathfrak{B}^2$，其中

$$\mathfrak{B}^1 \upharpoonright \mathscr{L}^0 \equiv_{\mathscr{L}^0_{AP}} \mathfrak{B}^2 \upharpoonright \mathscr{L}^0.$$

据 Robinson 一致性，$\varphi^1 \wedge \neg\varphi^2$ 有一概率模型，而这与$\models \varphi^1 \to \varphi^2$ 矛盾．证毕

对于无穷概率逻辑 $\mathscr{L}_{AP}$ 及其模型论的研究已经取得许多重要的成果，但也还留下许多问题．Keisler 认为至少有下列问题需要研究和解决．

（1）建立 $\mathscr{L}_{AP}$ 的带全称量词(x)的逻辑形式．

（2）建立一种逻辑有全称量词(x)和内外测度至少r的量词．

（3）研究诸如 $\mathscr{L}_{AP}$ 的逻辑使得适于它的模型有无穷测度而不是概率测度．

（4）研究诸如 $\mathscr{L}_{AP}$ 的逻辑使得适于它的模型有两个测

度(以及对应的量词). 得到适于两个测度 μ, ν 的模型的完全性定理, 其中 (i) μ 正交于 ν, (ii) μ 相对 ν 绝对连续.

(5) 从本章考察的逻辑观点来研究概率模型上的各种运算.

(6) 当 $\mathscr{L}$ 有函数符号时, 能无困难地贯彻分级 $\mathscr{L}_{AP}$ 上的种种成果? 当 $\mathscr{L}$ 有函数符号时, $\mathscr{L}_{AP}$ 上的成果能贯彻吗?

(7) 用诸如 $\mathscr{L}_{AP}$ 的种种逻辑重新考察抽象模型论.

后　　记

本书主要关注的是现代归纳逻辑的产生和发展，而把概率逻辑作为它的伴生形式和发展形式来研究。

众所周知，现代归纳逻辑产生于本世纪20年代，到50年代初，由R. Carnap发扬光大。从50年代到60年代中期，G. H. von Wright和J. Hintikka曾以极大的兴趣研究了归纳逻辑并取得一些重要成果（von Wright的工作本书没有直接介绍，但他的一部分工作我们在第八章概率语义学中间接做了介绍，二元概率语义学的条件8.11—8.18主要来自von Wright的工作）。可以说50—60年代是现代归纳逻辑的鼎盛时期。到了70年代，虽然出现了Cohen、Burks的模态归纳逻辑以及其他的归纳逻辑理论，但总的来说，逻辑学家渐渐失去了对现代归纳逻辑的兴趣。这主要表现在没有著名的有成就的逻辑学家长期致力于对现代归纳逻辑的研究，由此在归纳逻辑上就没有什么重要的逻辑成果。现代归纳逻辑逐渐成为科学哲学家和其他领域的科学家经常或偶然光顾的场所(例如Cohen主要是科学哲学家，而Burks则是计算机学家)。而从那时起，演绎逻辑、特别是其中的哲学逻辑，却得到迅猛的发展，产生了模态逻辑、时态逻辑、道义逻辑等一大批成果极为丰富的新逻辑分支。直至今天，尽管还有人零星地发表一些关于归纳逻辑的文章，但总的来说，现代归纳逻辑是在当代逻辑学家，特别是大逻辑学家的视野之外。

现代归纳逻辑的衰落不是没有道理的。可以认为它是由以下两对矛盾造成的。首先现代归纳逻辑研究的对象是结论

范围超出前提范围的归纳推理，而它目前用来研究的形式化工具是经典一阶逻辑及其扩张、经典概率论和现代测度论.这样，用演绎逻辑和演绎推理来刻划本质上是非演绎的归纳推理,这本身不能不是一个矛盾. 其次,归纳逻辑学家想用自己构造的系统合理地、全面地刻划我们在经验科学和日常生活中使用的所有归纳推理,至少是主要的归纳推理,而归纳推理从现有的认识看来绝不是纯逻辑的东西，传统归纳推理的有效性(归纳逻辑学家称之为归纳法的正当性)决不能只用逻辑(通常理解为只与语言相关而与外部世界无关) 本身来说明，而是涉及到许多哲学问题. 这里有本体论问题(例如自然齐一性原则、无差别原则),也有认识论问题(例如合理相信度和实际相信度). 如果说第一对矛盾是用演绎推理刻划非演绎推理的矛盾，那么第二对矛盾就是用逻辑来解决哲学问题的矛盾，我们认为是第二对矛盾更为本质地导致了现代归纳逻辑的衰落. D. Hume 和 K. Popper 对古典归纳逻辑和现代归纳逻辑的批判不是没有道理的，他们也看到了由于上述矛盾的存在，使得归纳逻辑的有效性不可能奠定在纯逻辑的基础上. 归纳逻辑没有在根本上解决归纳问题. Carnap 想用纯逻辑的方法构造一个大一统的归纳逻辑，用以刻划所有常见的归纳推理，这本身包含用逻辑去代替或解决哲学问题的思想,这就不能不使他的逻辑变得苍白无力. 而象 Cohen 和Burks 那样建立的逻辑，旨在说明实际科学实验中推理的程序,由于他们没有处理好逻辑和(科学)哲学的矛盾,所以他们建立的语义学(严格说只是某种语义解释)概括性不强，过分拘泥于人们在实践中的推理程式. 这就使他们的逻辑杂乱,没有说服力,至少对逻辑学家来说是如此. 从他们的得失我们可以看到逻辑不能迁就哲学，而是应该从哲学中汲取营养,然后迅速地脱离哲学,沿着逻辑本身的要求发展.

我们认为现代归纳逻辑今后可能向两个主要方向发展。一个方向是将它作为科学方法论和科学哲学的问题看待。这时的归纳逻辑应该称为归纳理论或归纳哲学更好些。在这里不能把归纳法只作为形式推理来研究，而是要深入探讨归纳法与自然界的本性，与人的认知心理的关系。这个方向的工作主要应该由哲学家去做。 F. P. Ramsey, L. J. Savage, I. Levi 和 H. E. Jr. Kyburg 等人的工作应该看作是这个方向的工作。

第二个方向，主要是逻辑学家应该去做的工作，就是要建立有形式化语义学和语法系统的现代归纳逻辑（象 Adams 的条件句概率逻辑，象 Keisler 的无穷概率逻辑），只要它刻划了归纳推理的某个本质特性（例如归纳推理的概然性、不确定性）就行，而不去苛求这样的系统能否合理地全面地说明人们在实践中运用的所有归纳推理，不去苛求它们能否最终解决哲学上的归纳问题。就象对模态逻辑的 Kripke 语义学一样。现代模态逻辑学家并不为这种语义模型导致在某些情况下$\Box\varphi \rightarrow \varphi$不是有效式（历史上有许多逻辑学家认为$\Box\varphi \rightarrow \varphi$是有效式是天经地义的）而惶惶不可终日，他们并不为此抛弃相当优雅（简洁又而概括力强，由此既能建立许多有完全性定理的语法系统，又能讨论没有完全性定理的诸语法系统的本质特性）的 Kripke 语义学。（Carnap 后来对这种相对化也有所认识，原来他认为唯有他的 c^* 系统才是归纳逻辑系统，后来他认识到归纳逻辑系统有连续统多个，包括许多在常人看来行为古怪反常的归纳逻辑系统。）

从广义上（在概然推理的意义上，在不确定推理的意义上）说，现在发展得比较完善的无穷概率逻辑、多值逻辑（特别是连续值的多值逻辑），非单调逻辑以及与模态逻辑、时态逻辑、认知逻辑相结合的概率逻辑等都应该属于现代归纳逻辑

范畴，因为它们都在某种程度上、在某些侧面刻划了归纳推理的特性。

从 Kripke 语义学产生的相对化运动应该为现代归纳逻辑所接受。古老的归纳问题应该被分解。哲学家仍然可以根据每个时代所得到的知识去讨论归纳问题，而逻辑学家则应该根据新的逻辑思想和逻辑手段去建立各种逻辑系统，来分别刻划归纳法的某些特性，虽然这样做也许不可能最终解决归纳问题。

参考文献

[1] J. M. Keynes, A Treatise on Probability, McMillan, London, 1921.

[2] H. Reichenbach, The Theory of Probability, The University of California Press, Berkeley, 1949.

[3] R. Carnap, Logical Foundations of Probability, The University of Chicago Press, Chicago, 1951.

[4] R. Carnap, The Continuum of Inductive Methods, The University of Chicago Press, Chicago, 1952.

[5] R. Carnap and R.C. Jeffrey (eds), Studies in Inductive Logic and Probability, Vol.1, The University of California Press, Berkeley, 1971.

[6] J. Hintikka, A Two-Dimensional Continuum of Inductive Methods, Aspects of Inductive Logic, North-Holland Publ. Co., Amsterdam, 1966.

[7] I. Niiniluoto, On a K-Dimensional System of Inductive Logic, Philosophy of Science Association, 1976 (eds. by F. Suppe and P. Asquith), 1977.

[8] J. Hintikka, Towards a Theory of Inductive Generalization, Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science, North-Holland Publ. Co., Amsterdam, 1965.

[9] J. Hintikka, On a Combined System of Inductive Logic, Studia Logico-Mathematica et Philosophia in Honorem Rolf Nevanlinna, Acta Philosophica Fennica, Vol. 18, 1965.

[10] J. Hintikka and I. Niiniluoto, Axiomatic Foundation for the Theory of Inductive Generalization, Formal Methods in the Methodology of the Empirical

Sciences, D. Reidel, Dordrecht, 1976.

[11] J. Hintikka, Distributive Normal Forms in First-Order Logic, Formal Systems and Recursive Functions, North-Holland Publ. Co., Amsterdam, 1965.

[12] E.W. Adams, Probability and the Logic of Conditionals, Aspects of Inductive Logic, North-Holland Publ. Co., Amsterdam, 1966.

[13] L.J. Cohen, The Probable and the Provable, Clarendon Press, Oxford, 1977.

[14] L.J. Cohen, Comments and Replies, Applications of Inductive Logic, (eds. L. J. cohen and Mary Hesses,), Clarenden Press. Oxford. 1980, p.65.

[15] L.J. Cohen, The Implication of Induction, Methuen & C. Ltd, London, 1970.

[16] A. W. Burks, Chance, Cause, Reason, The University of Chicago Press, Chicago, 1977.

[17] H. Leblanc, Alternatives to Standard First-order Semantics, Handbook of Philosophical Logic, Volume I, (eds, by D. Gabbay & F. Guenthner), D. Reidel, Dordrecht, 1983.

[18] D. Scott & P. Krauss, Assigning Probabilities to Logical Formulas, North-Hocland Publ. Co., Amsterdam, 1966.

[19] D. N. Hoover, Model Theory of Probability Logic, Doctoral Dissertation, University of Wisconsin, 1978.

[20] H. J. Keisler, Hyperfinite Model Theory, Logic Colloquium' **76**, (eds. J. M. E. Hyland and R. O. Gandy), North-Holland Publ. Co., 1977.

[21] H. J. Keisler, Probability Quantifiers, Model-Theoretic Logic, Springer-Verlag, Heidelberg, 1985.

[22] H. J. Keisler, Model Theory for Infinitary Logic North-Holland Publ. Co., Amsterdam, 1971.